UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

INSTITUTO DE ECOLOGÍA

BIOLOGÍA EVOLUTIVA

Variación fenotípica y predicción genómica en *Abies religiosa* (Pinaceae)

TESIS

PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS BIOLÓGICAS

PRESENTA:

M. Sc. SEBASTIÁN ARENAS JIMÉNEZ

TUTOR PRINCIPAL: DR. JUAN PABLO JARAMILLO CORREA

INSTITUTO DE ECOLOGÍA, UNAM

COMITÉ TUTOR: DR. VÍCTOR LUIS BARRADAS MIRANDA

INSTITUTO DE ECOLOGÍA, UNAM

COMITÉ TUTOR: DR. JORGE NIETO SOTELO

JARDÍN BOTÁNICO, UNAM

CIUDAD DE MÉXICO, ABRIL, 2021

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS BIOLÓGICAS

INSTITUTO DE ECOLOGÍA

BIOLOGÍA EVOLUTIVA

Variación fenotípica y predicción genómica en *Abies religiosa* (Pinaceae)

TESIS

PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS BIOLÓGICAS

PRESENTA:

M. Sc. SEBASTIÁN ARENAS JIMÉNEZ

TUTOR PRINCIPAL: DR. JUAN PABLO JARAMILLO CORREA

INSTITUTO DE ECOLOGÍA, UNAM

COMITÉ TUTOR: DR. VÍCTOR LUIS BARRADAS MIRANDA

INSTITUTO DE ECOLOGÍA, UNAM

COMITÉ TUTOR: DR. JORGE NIETO SOTELO

JARDÍN BOTÁNICO, UNAM

CIUDAD DE MÉXICO, ABRIL, 2021

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS
ENTIDAD INSTITUTO DE ECOLOGÍA
OFICIO CPCB/354/2021
ASUNTO: Oficio de Jurado

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
P r e s e n t e

Me permito informar a usted que en la Sesión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día **08 de marzo de 2021,** aprobó el siguiente jurado para la presentación de examen para obtener el grado de **DOCTOR EN CIENCIAS,** del estudiante **ARENAS JIMÉNEZ SEBASTIÁN** con número de cuenta: **513453108,** con la tesis titulada: **"Variación Fenotípica y Predicción Genómica en Abies religiosa (Pinaceae)",** bajo la dirección del **DR. JUAN PABLO JARAMILLO CORREA,** quedando integrado de la siguiente manera:

Presidente:	DR. JUAN SERVANDO NÚÑEZ FARFÁN
Vocal:	DR. ANTONIO GONZÁLEZ RODRÍGUEZ
Vocal:	DRA. ELLA GLORIA VÁZQUEZ DOMÍNGUEZ
Vocal:	DRA. ANGELINA MARTÍNEZ YRIZAR
Secretario:	DR. JORGE NIETO SOTELO

Sin otro particular, me es grato enviarle un cordial saludo.

A T E N T A M E N T E
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Cd. Universitaria, Cd. Mx., a 27 de abril de 2021

COORDINADOR DEL PROGRAMA

POSGRADO EN CIENCIAS BIOLOGICAS
COORDINACIÓN

DR. ADOLFO GERARDO NAVARRO SIGÜENZA

AGRADECIMIENTOS

Inicialmente, quisiera agradecer a la Universidad Nacional Autónoma de México y al programa de Doctorado en Ciencias Biológicas, por la formación académica y personal que diariamente me ha dado, y por haberme dado un espacio durante el tiempo del doctorado. Especialmente a su personal administrativo, académico y técnico por las facilidades otorgadas durante mis estudios.

El mismo agradecimiento le ofrezco al Consejo Nacional de Ciencia y Tecnología (CONACyT), ya que fue el sustento financiero de tesis con diferentes proyectos (CB-2016-284457, 278987, CB0153305) y de mi persona (480152). De igual manera, estoy muy agradecido con el programa de financiamiento PAEP que me permitió asistir a un congreso internacional (en Carolina del Norte, Estados Unidos).

A la Dirección General de Asuntos del Personal Académico de la UNAM, DGAPA; PAPIIT proyecto IN208416 asignado al Dr. Juan Pablo Jaramillo Correa.

De forma especial agradezco los comentarios, sugerencias recomendaciones y consejos de mi tutor, Dr. Juan Pablo Jaramillo Correa.

En el mismo sentido, estoy muy agradecido con los miembros de mi comité tutoral (Drs. Juan Pablo Jaramillo Correa, Víctor Barradas Miranda y Jorge Nieto Sotelo) por todas las discusiones que tuve con ellos para enfocar y desarrollar mejor mi proyecto y por todo el apoyo que siempre he recibido de su parte.

AGRADECIMIENTOS PERSONALES

Inicialmente agradezco también al Instituto de Ecología de la Universidad Nacional Autónoma de México que ha sido mi casa durante 6 años de mi vida, siempre he sido y seguiré siendo un puma.

Ofrezco mi gratitud individual e infinita al Dr. Juan Pablo Jaramillo, porque desde antes de entrar al programa de Doctorado, siempre estuvo dispuesto a hablar conmigo sobre mis ideas y fomentó gran parte de mi curiosidad, creatividad y hambre de conocimiento, gracias por aceptarme en el proyecto y por el gran apoyo durante todo el Doctorado. ¡Espero sin duda que continuemos colaborando!

Agradezco de forma infinita a los honorables miembros del jurado Drs. Juan Núñez Farfán, Ella Gloria Vázquez Domínguez, Antonio González Rodríguez, Jorge Nieto Sotelo y Angelina Martínez Yrizar por haber leído y haber aportado ideas claves para mejorar esta tesis.

Estoy eternamente agradecido con el Consejo Civil Mexicano para la Silvicultura Sostenible (CCMSS) por su apoyo incondicional, particularmente a la Dra. Lucía Madrid y al ingeniero Andrés Juárez, junto con los señores ejidatarios de Amanalco de Becerra (Estado de México) que me abrieron las puertas de sus rodales forestales para permitirme así, llevar a cabo la investigación. De igual Igualmente, estoy muy agradecido con el Dr. Andrés Cortés y la Dra. Alicia Mastretta por el apoyo, amistad y enseñanzas que me dieron, y por las discusiones y colaboraciones que hemos mantenido. Sin duda, parte de mi madurez científica se debe a todo lo que me han enseñado y por su exigencia de que todo quede lo más claro posible. Agradezco enormemente a Brian Boyle y su equipo, por haberme secuenciado con la tecnología GBS nuestras muestras de oyamel.

Igualmente conservo gran gratitud con mi jurado de Candidatura al título de Doctor en Ciencias (Drs. Ana Weiger, Luis David Alcaraz, Antonio González, Jorge Nieto y Rocío Cruz) por todos sus aportes para mejorar el desarrollo y la logística de esta tesis. Agradezco a la Dra. Tania Garrido, así como a Anabel por todo el apoyo técnico que me dieron en el laboratorio. También Agradezco a los Drs. Daniel Piñero, Ella Vázquez y María del Carmen Madujano por haberme aceptado en su laboratorio hace ya cuatro años.

Este trabajo no habría sido posible sin la ayuda de Ernesto Campos Murillo por permitirme el acceso al *cluster* de la CONABIO para el proceso bioinformático de datos, de Gustavo Giles

por su apoyo en el ensamble de los datos, de Leopoldo Vázquez por su apoyo en los análisis. De igual forma agradezco a mis compañeros de proyecto, Verónica Reyes, Jorge Cruz y nuevamente a Gustavo Giles; al igual que a Armando Sunny y Karen Carrasco por ayudarme con las colectas en campo, las extracciones y por las discusiones. ¡arriba el equipo oyamel!

Agradezco a todos los miembros actuales y pasados del Laboratorio de Genética y Ecología. Me siento muy satisfecho y contento de haber pasado este tiempo con todos ustedes y solo tengo una sensación de gratitud y cariño. Gracias a todos por apoyarme en el crecimiento profesional e intelectual, nunca olvidaré el Mal del Puerco, las posadas de fin de año, las rutinas del gimnasio con Madison y Alejandro, las discusiones de coníferas o las 4 paredes del cubículo que me acompañaron durante 4 años.

No me alcanzan las palabras para agradecerle a mis padres, Silvia Patricia y Ariel Eduardo para decirles lo afortunado que me siento de que sean parte de mi vida ¡Los quiero y los extraño entrañablemente¡ Con mucho cariño, agradezco a toda mi familia por todo su apoyo, fortaleza y tolerancia. Me siento muy afortunado tres hermosos hermanos, Juan Guillermo, Carlos Fernando y María Juliana, por mis dos muy queridos abuelos, Doña Silvia y Don Jaime (*Requiéscat in pace*) por su infinito apoyo y amor y por mis tíos incondicionales. Agradezco también a Lina Bolívar, por siempre sacar lo mejor de mí, acompañarme en cada decisión, y ayudarme a ser una mejor persona.

Finalmente, agradezco a mis amigos, Daniel, Christian, Víctor, Mariela, Gaby, Karol, Yury, Angie, Juanse, Sebas, David, Paola, Ana, Laura, Juliette, Susana, Wilson, Tati Colmenares, Tati Villamizar, Karla, Andrés, en fin, a todos mis amigos de la vida por todo lo que hemos convivido, reído, discutido, aprendido. Son una parte importante de mi vida ¡Los llevo en el alma!

Esta tesis es dedicada en primer lugar a mi abuelo quién me enseñó a amar la ciencia sin si quiera tener una carrera; abuelito, mi corazón te extraña, mi mente te recuerda y mi alma te echa de menos (RIP, 09/01/21). En segundo lugar, a todos los fallecidos por el COVID a causa de la falta de conciencia y tolerancia social en especial a Don Rafael Bolívar, me quedé con las ganas de estrechar su mano. Y en último lugar a mi pequeña y peluda compañía que me acompañó desde el día que decidí inscribirme en la licenciatura.

PREFACIO

La presente investigación es un trabajo original e independiente del autor, S. Arenas. Desarrollé las preguntas y objetivos de investigación a través de múltiples consultas con mi comité tutoral. Colaboré con colegas en las etapas de obtención de datos. Realicé el análisis de los mismos, la interpretación de los resultados y preparé los manuscritos. Diferentes secciones de esta tesis fueron o serán enviadas para publicación tal como se enumera a continuación:

CAPÍTULO II

• **Arenas, S**, Campo, J, Mastretta-Yanes, A, Jaramillo-Correa JP, 2021. Within-population genotype – soil interactions drive phenotypic variation in a recovering fir forest from central Mexico. Artículo aceptado en *Forest Ecology and Management*.

CAPÍTULO III

• **Arenas, S**., Jaramillo-Correa, J.P., 2020. ¿Es posible utilizar modelos de selección genómica (SG) en poblaciones naturales de plantas con largos tiempos generacionales y plantear perspectivas de manejo/conservación? Artículo en revisión en *Annals of Forest Science*

CAPÍTULO IV

• **Arenas, S.,** Cortés, A.J., Mastretta-Yanes, A., Jaramillo-Correa, J.P. 2021. Evaluating the accuracy of genomic prediction for the management and conservation of relictual natural tree populations. *Tree Genetics & Genomes* 17, 12. https://doi.org/10.1007/s11295-020-01489-1

TABLA DE CONTENIDO

RESUMEN

Comprender cómo responden los componentes de la variación fenotípica a la heterogeneidad del ambiente proporciona información importante para conservar y manejar las poblaciones naturales. La variación fenotípica es fundamental porque las diferencias entre individuos sirven como marcadores para estudiar los factores genéticos y ambientales responsables de ciertos rasgos específicos. Los árboles forestales, como el oyamel, *Abies religiosa* (Kunth) Schltdl. et Cham., dependen de sus niveles de variación genética para adaptarse a la heterogeneidad del ambiente. Por lo tanto, su cantidad y distribución en el espacio de la variación genética son importantes para la supervivencia y la adaptabilidad a largo plazo de las poblaciones. Las variaciones del ambiente afectan el fenotipo de los árboles a diferentes escalas geográficas y temporales. La variación del fenotipo está generalmente codificada por muchos genes que responden al ambiente, resultando en una interacción genotipo y ambiente (G × E). Esta interacción indica que los factores locales, como por ejemplo la variación edáfica dentro de una población, influye en el fenotipo de manera diferencial dependiendo del genotipo del individuo, lo que resulta en fenotipos contrastantes. Estudiar esta interacción e integrarla en modelos predictivos puede ayudar a mejorar programas de reforestación, manejo y conservación de especies, incluidos los *Abies*.

Los árboles forestales presentan múltiples desafíos para el estudio de su capacidad de adaptación al ambiente, como su madurez sexual tardía o el retraso en la expresión de fenotipos importantes relacionados con la productividad, lo que provoca ciclos de selección largos y recurrentes. La Predicción Genómica (PG) se está aplicando como una herramienta para abordar tales deficiencias mediante la predicción temprana de fenotipos, utilizando simultáneamente los rasgos comerciales y un extenso conjunto de marcadores moleculares distribuidos a lo largo del genoma, por lo general polimorfismos de nucleótido único (SNP). Aunque algunos de los estudios ya se han desarrollado en plantaciones comerciales, aún no se ha probado hasta qué punto esta herramienta podría ser útil para hacer predicciones en poblaciones forestales naturales. Esto permitiría manejar la diversidad genética e impulsar los estudios de adaptación al ambiente. Para abordar estos desafíos, estructuré esta tesis en tres análisis que utilizan varias metodologías, basándose en fenotipos de crecimiento y fisiología, y técnicas de genotipado de alto rendimiento en dos grupos de árboles de *A. religiosa* con orígenes diferentes (regeneración natural y

reforestación). El primer estudio explora modelos poligénicos microgeográficos para caracterizar la diferenciación genética, las presiones de la heterogeneidad del suelo y la interacción G × E sobre la variación cuantitativa. El segundo se centra en revisar las posibilidades teóricas para la implementación de herramientas de PG en poblaciones naturales, sugiriendo con ello, la preselección de individuos para reforestar y recuperar ecosistemas degradados. El tercer estudio investiga el uso de PG para estimar la precisión en las estimaciones de parámetros genéticos en poblaciones naturales de árboles.

Los resultados de estos tres estudios demostraron que: *i*) una parte de la diferenciación fenotípica podría explicarse por la interacción de factores genéticos, y edáficos dentro del sitio del ensayo (G × E), *ii*) existen genes potencialmente adaptativos que podrían estar respondiendo a escalas pequeñas dentro de un mismo rodal forestal, *iii*) la PG es una herramienta eficaz para orientar programas de manejo y conservación en poblaciones naturales para predecir fenotipos individuales y *iv*) los modelos de PG más eficientes fueron los construidos con árboles naturales y usados para predecir el desempeño de individuos reforestados en el mismo ambiente. Finalmente, se resaltan los factores que dificultan la transferencia de los modelos a las poblaciones naturales, como por ejemplo el tamaño de muestra para el fenotipado y genotipado y la variación en los rasgos entre individuos jóvenes y adultos. Además, se analizan algunas perspectivas para explorar estas metodologías en poblaciones naturales, a medida que las tecnologías de secuenciación de ADN continúan mejorando, y aumenta la calidad y cantidad de recursos genómicos disponibles para organismos con genomas complejos.

ABSTRACT

Understanding how the components of phenotypic variation respond to environmental heterogeneity provides fundamental information for conserving and managing natural populations. Phenotypic variation is fundamental because differences among individuals serve as markers to study the genetic and environmental factors responsible for specific traits. Forest trees, such as *Abies religiosa* (Kunth) Schltdl. et Cham, depend on their levels of genetic variation to adapt to environmental heterogeneity. Therefore, their number and distribution in space are important for the survival and long-term adaptability of populations. Environmental variations affect tree phenotypes at different geographic and temporal scales. Phenotype variation is generally encoded by many environment-responsive genes, resulting in a genotype-environment interaction (G × E). This interaction indicates that local factors, such as edaphic variation within a population, differentially influence depending on the genotype of the individual, resulting in contrasting phenotypes. Studying this interaction and integrating it into predictive models can help to improve reforestation, management and conservation programs for species, including *Abies*.

Forest trees present multiple challenges to study adaptive capacity to the environment, such as late sexual maturity or delayed expression of important phenotypes related to productivity, resulting in long and recurrent selection cycles. Genomic Prediction (GP) is being applied as a tool to address such deficiencies by early prediction of phenotypes using simultaneously commercial traits and a dense set of molecular markers distributed across the genome, usually single nucleotide polymorphisms (SNP). Although some of the studies have already been developed in commercial plantations, the extent to which this tool could be useful for making predictions in natural forest populations remains to be tested. This would allow us to manage genetic diversity and drive studies of adaptation to the environment. To address these challenges, I structured this thesis into three analyses using various methodologies, based on growth and physiology phenotypes and high-throughput genotyping techniques in two groups of *A. religiosa* trees with different origins (natural regeneration and reforestation). The first study explores microgeographic polygenic models to characterize genetic differentiation, the pressures of soil heterogeneity and G × E interaction on quantitative variation. The second focuses on reviewing the theoretical possibilities for the implementation of GP tools in natural populations, thereby

suggesting the pre-selection of individuals for reforestation and recovery of degraded ecosystems. The third study investigates the use of GP to estimate the precision of genetic parameter estimates in natural tree populations.

The results of these three studies showed that: (*i*) part of the phenotypic differentiation could be explained by the interaction of genetic, and edaphic factors within the trial site (G × E), (*ii*) there are potentially adaptive genes that could be responding at small scales within the same forest stand, (*iii*) GP is an effective tool to guide management and conservation programs in natural populations to predict individual phenotypes, and (*iv*) the most efficient GP models were those built with natural trees and used to predict the performance of reforested individuals in the same environment. Finally, factors that make it difficult to transfer the models to natural populations, such as sample size for phenotyping and genotyping and variation in traits between young and adult individuals, are highlighted. In addition, some prospects for exploring these methodologies in natural populations are discussed, as DNA sequencing technologies continue to improve, and the quality and quantity of genomic resources available for organisms with complex genomes increases.

CAPÍTULO I

INTRODUCCIÓN

Los árboles forestales constituyen aproximadamente el 80% de la biomasa de la tierra, almacenando una gran cantidad de carbono (Ellison et al., 2017). Cumplen funciones ecosistémicas integrales en el mantenimiento de la biodiversidad, la protección de los recursos hídricos y edáficos, y el secuestro de carbono; además, son un componente importante a nivel cultural y económico para las poblaciones humanas (Daniels, 1984; FAO, 2014; Le et al., 2012; Schwartz et al., 2012). Sin embargo, durante las últimas décadas los impactos antropogénicos (p. ej. la deforestación y la fragmentación del paisaje) han devastado bosques por todo el planeta (Isabel et al., 2020; Lobell y Gourdji, 2012). Además, se predice que el actual calentamiento climático (CC) acelerará el ritmo de los cambios ambientales, así como el aumento en la frecuencia de eventos de sequía e inundaciones, las infestaciones de plagas y las variaciones en la temperatura, entre otras comprometiendo la resiliencia de las poblaciones y el funcionamiento de los ecosistemas (Breshears et al., 2013; Hooper et al., 2012; Nadeau et al., 2016; Raffa et al., 2008). El éxito en la respuesta de los árboles a estos cambios dependerá de la capacidad de migrar a ambientes más favorables y de la variación genética disponible entre y dentro de las poblaciones/especies para hacerle frente a las condiciones locales cambiantes y adaptarse o responder plásticamente al ambiente (Aitken et al., 2008; Alberto et al., 2013; Hansen et al., 2012; Koskela et al., 2014; Plomion et al., 2016)

1.1 Variación genética en árboles

El estudio de la variación dentro y entre poblaciones/especies es un tema de constante análisis para los genetistas. En general, la variación genética es la base para el cambio evolutivo potencial, y se le considera el nivel básico de diversidad biológica (Fox y Wolf 2006; George et al., 2017; Tamaki et al., 2008). Esta diversidad es crucial para la aptitud y supervivencia de los individuos y la capacidad de adaptación a los cambios ambientales (Melosik et al., 2016; Holliday et al., 2017; Hamabata et al., 2019). La presencia de suficiente variación genética en los árboles (y en casi cualquier otra especie) es crucial para la persistencia de las poblaciones. La pérdida de variación genética afectará la capacidad de respuesta por parte de las poblaciones a la variación ambiental, la pérdida del hábitat y la presencia de nuevos patógenos (Aitken et al., 2008; Isabel et al., 2020; Oddou-Muratorio et al., 2020). Por ejemplo, el estrés hídrico es una presión selectiva importante, y la variación genotípica en la resistencia a la deshidratación en árboles (Muthoo, 2002; Soltys-Kalina et al., 2016) resulta en la supervivencia diferencial de algunos de ellos

después de un evento de sequía (Moran et al., 2017; Sallam et al., 2019). Estudios recientes han analizado los patrones geográficos de la variación genética en especies forestales y cómo éstos afectan su capacidad para sobrellevar el CC (p. ej. Carvalho et al., 2020; Collevatti et al., 2019). Estos estudios, mediante evaluaciones de correlación entre la variación genética y variables ambientales, documentan que las poblaciones están adaptadas a las condiciones locales. Por lo tanto, las diferencias regionales con respecto al impacto del CC y la evolución del nicho ecológico conducirán a eventos maladaptativos para las poblaciones, y algunas de éstas se enfrentarán a una mayor probabilidad de extinción que otras (Koskela et al., 2014; Schneider et al., 2011).

Muchas especies forestales presentan altos niveles de diversidad genética, lo que en principio les brinda la capacidad de abarcar grandes áreas de distribución y adaptarse a entornos fluctuantes (Balvanera y Aguirre, 2006; Kremer et al., 2014; Savolainen et al., 2007; Sork, 2018). Si se desea poder predecir con precisión el impacto de las variaciones del ambiente, es necesario comprender cómo se distribuye la variación genética dentro y entre especies/poblaciones, incluso a escalas geográficas finas (i.e. pequeñas áreas geográficas; Kubota et al., 2015; Rae, 2013). Por tanto, es fundamental evaluar la diversidad genética existente y la variación fenotípica vinculada a ésta, para explicar cómo las especies sobreviven y se adaptan a las presiones selectivas locales (Blanquart et al., 2013; Lascoux et al., 2016; Sork, 2018). Sin embargo, predecir la arquitectura genética de la adaptación en poblaciones naturales es difícil; al menos sin tener un enfoque generalizado para evaluar la sensibilidad a ambientes heterogéneos (Housset et al., 2018; Merilä y Hendry, 2014).

1.2 Genética cuantitativa

La variación fenotípica es fundamental para entender los procesos evolutivos, ya que no solo es sobre los fenotipos que actúa la selección natural, sino que además las diferencias entre individuos sirven como marcadores para estudiar los factores genéticos y ambientales responsables de rasgos específicos (p. ej. anatómicos, morfológicos o funcionales; Fox y Wolf 2006). Hay dos tipos de variación basada en estos rasgos, la cualitativa y la cuantitativa. El rasgo cualitativo es aquel en el que hay un número de fenotipos de tipo discreto o categórico. Generalmente, un número reducido de genes participan en el control de tales rasgos, por eso, se ven en menor medida afectados por el medio ambiente (Kahlke y Hon, 2014). Los ejemplos de rasgos cualitativos incluyen: la presencia o ausencia de enfermedades hereditarias o congénitas,

la forma del fruto, el color de las estructuras florales, la ruta fotosintética, entre otros (Li et al., 2017). El rasgo cuantitativo es aquel para el que existen fenotipos que no pueden clasificarse fácilmente en categorías discretas (Grattapaglia et al., 2018; Ikram y Chardon, 2010). En estos últimos, un gran número de genes o *loci* con un efecto pequeño y aditivo influyen en la variación del rasgo (poligénicos); por ello, son útiles para comprender la mecánica y las bases genéticas que orientan los procesos de adaptación local (Alberto et al., 2013; Housset et al., 2018). Tradicionalmente, la adaptación local se ha descrito en experimentos de jardín común con germoplasma de diferentes procedencias plantado en las mismas condiciones ambientales; en ellos se evalúan rasgos relacionados al *fitness* como el crecimiento, la morfología y la supervivencia (Blanquart et al., 2013; Lascoux et al., 2016; Valladares et al., 2014).

1.3 Variación fenotípica y heterogeneidad ambiental

Las variaciones del ambiente influyen en la expresión del fenotipo de los árboles a nivel intraespecífico y a diferentes escalas geográficas. A gran escala (macroambiente), los fenotipos responden a clinas ambientales de elevación, temperatura y/o precipitación, así como a la variación en la composición de suelo (De Mita et al., 2013; de Villemereuil et al., 2016; Zhang et al., 2019). A escala fina (microambiente), la variación fenotípica entre árboles cercanos dentro un mismo rodal es causada principalmente por variaciones edáficas y lumínicas, e interacciones bióticas, como la presencia de plantas nodrizas, comunidades de microorganismos y/o la exposición a plagas (Cappai et al., 2017; Carbajal-Navarro et al., 2019; Guerrero et al., 2018). Así, las variables del macro y microambiente pueden interactuar con la cantidad de variación genética (y el genotipo mismo) de los individuos, afectando directamente su arquitectura (Finkeldey y Hattemer 2007; Des Marais et al. 2013). Sin embargo, en la mayoría de los estudios realizados en especies forestales solo se ha evaluado la variación fenotípica a nivel de población en respuesta al componente genético y ambiental a escala geográfica grande (estimando heredabilidades y correlaciones genéticas) (de los Campos et al., 2015; Hodge y Dvorak, 2015), dejando rezagada la variación a nivel microgeográfico. Aún hace falta evaluar si esta variación microambiental es relevante en los rasgos relacionados con el *fitness* o adecuación biológica de la población.

1.4 El suelo como factor selectivo microgeográfico

El suelo es el principal almacén de carbono (C) y nitrógeno (N) en formas disponibles para mantener las funciones de los organismos asociados (p. ej., plantas y micorrizas); por lo que actúa como una presión selectiva sobre ellos. La dinámica del C en el suelo es el resultado del balance entre la fotosíntesis (fijación de C) y la respiración (mineralización del C; Lafleur et al., 2015). El CO_2 atmosférico es incorporado a los ecosistemas por medio de la fotosíntesis de los organismos autótrofos, como las plantas con clorofila, y los microorganismos quimioautótrofos, que convierten el CO_2 a carbohidratos para integrarlos a su biomasa. Con la muerte, la biomasa se deposita en el suelo en forma de residuos orgánicos, donde la fauna y los microorganismos del suelo descomponen los residuos vegetales, que pasan a formar parte de la materia orgánica (MO). En el suelo, la MO formada por partículas con diferentes niveles de descomposición, así como por compuestos macro y micromoleculares, representan una fuente de energía accesible al microbioma (Silfver et al., 2015).

Así como el C, el N forma parte importante del ciclo biogeoquímico del ecosistema, ya que es un elemento esencial para la vida, formando parte del ADN y las proteínas (Chen et al., 2019; Elliot et al., 2019). El N atmosférico (N_2) se fija mediante procesos biológicos, por medio de bacterias (como *Rhizobium* y *Bradyrhizobium*) que cuentan con las enzimas necesarias para reducir el N_2 a formas reactivas como el amonio (NH_4^+), un compuesto biológicamente asimilable (Pugnaire et al., 2019; Sauer et al., 2012). Dentro de la solución del suelo, los microorganismos y plantas compiten por el NH_4^+, ya que esta es la forma de N preferida para el metabolismo y para la síntesis de proteínas (Nacry et al., 2013; Potter and Snyder, 1915; Sauer et al., 2012). Cuando las plantas adquieren el NH_4^+, incorporan principalmente el N para el crecimiento. Por lo tanto, la varianza en el crecimiento de las plantas de alguna manera refleja la disponibilidad edáfica de C, MO, N_2 y NH_4^+ (Madritch et al., 2009; Raven and Andrews, 2010).

Los procesos ecosistémicos orientados por la variabilidad de rasgos poligénicos de los árboles, como el crecimiento y la tasa fotosintética a escala fina, pueden ser regulados por la interacción entre el genotipo de la planta, la fertilidad del suelo y la estructura de las comunidades asociadas a ésta varianza fenotípica (Guerrero et al., 2018; Schweitzer et al., 2011). Por lo tanto, la variación del fenotipo es impactada por muchos genes que responden a la disponibilidad microambiental de nutrientes (Chen et al., 2019; Kubota et al., 2015), resultando en una interacción genotipo y ambiente (G × E), que implica algo más que la simple aditividad de ambos

componentes (Charmantier, 2014; Finkeldey y Hattemer 2007). Las interacciones G × E indican que los factores abióticos locales, como la variación espacial nutricional dentro del sitio (Cappai et al., 2017; Liu et al., 2019), la adición de fertilizantes (Bruelheide et al., 2018; Le et al., 2012), y la variación genética (Bailey et al., 2009; Madritch et al., 2006; Pregitzer et al., 2013), pueden influir potencialmente en la expresión de fenotipos adaptativos y, a su vez, tener impacto en la comunidad y en los procesos del ecosistema.

Diversas metodologías se han utilizado para analizar la interacción G × E y describir la distribución de sus componentes en combinación con análisis de genética cuantitativa. Con ellos se desarrollan modelos predictivos utilizando información poligénica; como los análisis de correspondencia canónica (CCA), de componentes principales (PCA), de redundancia (RDA) o los modelos lineales (Forester et al., 2018; Rellstab et al., 2015; Scotti et al., 2016). Estos estudios demuestran que cuando se incluye la información de cientos a miles de genes se pueden obtener buenos estimados de la varianza de los rasgos y su correlación con los factores ambientales, además de diseccionar los patrones de G × E. En el capítulo II de esta tesis se utilizan métodos de análisis de la adaptación poligénica para describir los componentes de la variación fenotípica en respuesta a la heterogeneidad en las propiedades físicas y químicas del suelo a escala fina dentro de un rodal forestal. Cabe anotar que simultáneamente se han desarrollado métodos más complejos y estadísticamente más sólidos (p. ej., predicciones bayesianas) para efectuar predicciones más precisas y confiables de la varianza fenotípica (Scotti et al. 2016) y algunos se exploran dentro de este trabajo.

1.5 Desafíos para estudiar las especies forestales y los marcadores moleculares

La capacidad de las poblaciones para adaptarse a la heterogeneidad ambiental va a depender de la cantidad de variación genética presente en los genes adaptativos (Aitken et al., 2008; Holliday et al., 2017; Sork, 2018). No obstante, los árboles forestales presentan múltiples desafíos para estudiar dicha variación, como se ha observado en el diseño de programas reproductivos y en ciertos experimentos (p. ej. al buscar *loci* asociados a rasgos cuantitativos (QTL), efectuar mapas de ligamiento y secuenciar genomas completos), debido a sus largos tiempos generacionales. Recientemente, con la llegada de las técnicas de marcadores moleculares basados en PCR (p. ej. microsatélites), bibliotecas de cADN y los avances en la secuenciación masiva o de nueva generación (p. ej., para generar miles de polimorfismos de un único nucleótido o SNPs),

han facilitado las investigaciones en organismos no modelo (Neale y Kremer, 2011; Savolainen et al., 2013). Recientemente, el uso de SNPs se ha hecho cada vez más frecuente; representan sustituciones de una sola base en una ubicación particular a lo largo del genoma; se encuentran con frecuencia en genes codificantes y, a menudo, son bialélicos, lo que los hace útiles para efectuar perfiles de variación genética dentro y entre poblaciones (Black et al., 2001; Yousefi et al., 2018).

1.5.1 Mapeo de loci de rasgos cuantitativos y estudios de asociación genética

El desarrollo de tecnologías rentables para obtener marcadores moleculares ha facilitado el estudio y la descripción de la arquitectura genética de la adaptación local, vinculando la información fenotípica con los genes o alelos subyacentes (neutrales o adaptativos). Así, el mapeo de QTL y los estudios de asociación genética (GWAS) se han desarrollado para incrementar la comprensión de la base genética de los rasgos cuantitativos (White et al. 2007). El mapeo de QTL tiene como objetivo identificar regiones génicas específicas que tengan una fuerte influencia en el rasgo de estudio, y pueden ayudar a hacer predicciones fenotípicas sobre dicho rasgo. Por otro lado, los análisis de asociación de todo el genoma (GWAS) se utilizan para medir y analizar las variantes genéticas y asociarlas directamente con un fenotipo y estimar su efecto en el mismo, todo independientemente de la complejidad genética y de las interacciones que regulen su variación (Alimi, 2016; Korte y Farlow, 2013).

Los GWAS utilizan un gran número de SNPs obtenidos con técnicas de secuenciación masiva, como GBS y DArTSeq (Soto-Cerda y Cloutier, 2010; Varshney et al., 2015) para aplicar una serie de pruebas estadísticas en las que los SNPs son tratados como eventos independientes; estos luego se correlacionan al fenotipo mediante modelos lineales generalizados (GLM, *General Linear Model*) (Bush and Moore, 2012; Corvin et al., 2010). Generalmente, los rasgos complejos están regulados por muchos genes con efecto diferencial (Francia et al., 2005), por lo que actualmente se han implementado modelos que retienen el efecto de miles de genes (incluso aquellos con uno muy pequeño) para explicar la arquitectura genética de estos rasgos en ramas de la ciencia como la medicina o el mejoramiento de plantas y animales.

1.6 Estrategias clásicas de mejora de árboles

Los programas de mejoramiento en árboles con selección recurrente comenzaron aproximadamente en la década de los 50's, para especies forestales con importancia económica (Badenes et al., 2016; Isik, 2014; Iwata et al., 2016). Desde entonces se han desarrollado rápidamente, en parte siguiendo los avances de los sistemas de mejoramiento en cultivos y animales de granja. El objetivo de los programas de mejora en árboles es evaluar la arquitectura de los rasgos de interés y aumentar la frecuencia de los alelos benéficos (Sniezko y Koch, 2017; White et al. 2007), todo a partir de diseños familiares con cruzas controladas, en experimentos aleatorios y replicados. Los progenitores seleccionados son aquellos que tienen y permiten obtener y "mejorar" el fenotipo deseado.

1.6.1 Selección asistida por marcadores y predicción genómica

En el pasado, la selección asistida por marcadores (SAM) era considerada una estrategia para tomar decisiones de mejoramiento genético en árboles (Grattapaglia, 2014; Neale y Williams, 1991), que hacía uso del desequilibrio de ligamiento (LD) entre diferentes *loci* asociados a rasgos cuantitativos (QTL) y los marcadores genéticos utilizados (White et al., 2007). La estrategia SAM es un método para disminuir los largos ciclos reproductivos en árboles (Isik, 2014) y comprender mejor los rasgos complejos de sus fenotipos (Grattapaglia y Resende, 2011; Iwata et al., 2016). Esta estrategia supone que pocos marcadores con un gran efecto proporcionan información suficiente para predecir los fenotipos. Sin embargo, en árboles forestales, el SAM no presentó los resultados esperados, a causa de limitaciones como la naturaleza extremadamente poligénica de los rasgos de importancia económica, así como al hecho de que no se pueden aplicar en poblaciones diferentes a las estudiadas, a la presencia de interacciones G × E en la expresión de los QTLs y a un reducido desequilibrio de ligamiento (LD) entre marcadores y QTLs (Nakaya y Isobe, 2012; Poland y Rutkoski, 2016). Estas limitaciones se están superando parcialmente con el desarrollo de las metodología de selección y predicción genómica (Badenes et al., 2016).

La predicción genómica (PG) utiliza simultáneamente los fenotipos y de miles a cientos de miles de marcadores, que se analizan sin información *a priori* de su efecto en el fenotipo (Hayes et al., 2009; Meuwissen et al., 2001). La PG es capaz de capturar gran cantidad de variación fenotípica de rasgos cuantitativos, ya que presupone que al menos algunos de los muchos marcadores se encontrarán en LD con los QTLs del rasgo (Resende et al., 2012a). El enfoque PG

combina información fenotípica y genotípica en un conjunto de entrenamiento (TRN), para desarrollar modelos que predigan valores reproductivos genómicos (GEBV) en un grupo de candidatos a ser seleccionados o de validación (TST), para los que solo se requiere su información genotípica (Goddard y Hayes, 2009; Lin et al., 2014; Van Eenennaam et al., 2014).

Por otro lado, la selección genómica (SG) implica que los modelos de PG sean validados en la siguiente generación, pero esto solo ha sido probado en plantaciones de especies modelo (p. ej. *Eucaliptus globulus*, *Populus trichocarpa*, *Pinus taeda*, *Picea glauca*) luego de 1-2 ciclos de selección (Grattapaglia et al., 2018; Isik, 2014). El método de SG evita la necesidad de realizar las largas fases de prueba que los árboles requieren para obtener datos fenotípicos precisos, lo que resulta en una mayor ganancia genética por unidad de tiempo (Grattapaglia y Resende 2011; Grattapaglia, 2014). Esto ya se ha validado en la cría selectiva de animales y plantas de cultivos (Crossa et al., 2017; Meuwissen et al., 2016). Además, la implementación de una matriz de relaciones genéticas (*G*; VanRaden 2008) dentro de los modelos de PG facilita su montaje, ya que no se necesitan cruzas controladas; además, ofrece estimaciones más precisas, al considerar toda la varianza genética del conjunto TRN (Heffner et al. 2009). Queda por probar hasta qué punto esta herramienta podría ser útil para manejar la diversidad genética en poblaciones forestales naturales, donde puede haber un amplio espectro de relaciones familiares (presencia de hermanos completos y medios, relaciones históricas, etc.) (Charmantier et al., 2014).

1.6.2 Predicción y selección genómica en árboles forestales

El potencial de la PG y la SG ha sido explorado con datos empíricos en pocas especies de árboles forestales, todas establecidas en plantaciones y con diseños familiares conocidos. Algunas de las especies estudiadas son: *Eucalyptus* spp. (Ballesta et al., 2018, 2020; Müller et al., 2017 Suontama et al., 2019; Tan et al., 2017), *Pinus taeda* L. (Resende et al., 2012a; Zapata-Valenzuela et al., 2013), *Picea glauca* (Beaulieu et al., 2014a, 2014b), *Picea mariana* (Lenz et al., 2017), *Pinus pinaster* (Bartholomé et al., 2016; Isik et al., 2016), *Pseudotsuga menziesii* (Ratcliffe et al., 2019; Thistlethwaite et al., 2017, 2019, 2020), *Picea abies* (Chen et al., 2018; Lenz et al., 2020) y *Shorea platyclados* (Sawitri et al., 2020). En el capítulo III, se presenta información relevante sobre varias de estas investigaciones previas. Esta tesis comprende uno de los primeros intentos para evaluar la factibilidad y precisión de la PG en rodales naturales de *Abies religiosa* (capítulo IV).

1.7 Utilización de la PG en las poblaciones naturales forestales

La importancia y justificación de las investigaciones forestales en diferentes campos, incluida la PG, tendrán mayor prioridad hacia el futuro (Isik, 2014), dada la enorme variedad de bienes y servicios que proporcionan los bosques. Conservar los recursos genéticos forestales es por lo tanto vital, ya que los bosques son fundamentales para el crecimiento económico sostenible y la adaptación ambiental hacia el futuro (FAO, 2014). Las especies forestales pueden sobrevivir en una amplia gama de condiciones ecológicas, ya que han evolucionado durante períodos de grandes cambios climáticos; por esta razón, estudiar su variación genética en poblaciones silvestres es necesario para abordar el desafío de mitigar o adaptarse a futuros cambios ambientales (Aitken et al., 2008; Aitken y Bemmels, 2016).

Históricamente, la forma más común de evaluar los recursos genéticos son los ensayos de procedencias bajo diferentes condiciones ambientales (Koskela et al. 2014). El objetivo principal de estas investigaciones es identificar poblaciones de árboles con buen desempeño fenotípico y que estén suficientemente adaptadas para ser fuente de semillas para la reforestación o la transferencia de germoplasma (Aitken y Whitlock 2013; Lin et al., 2018; Ledig y Kitzmiller 1992). Dado que los ensayos son costosos de establecer y mantener, y al largo período que toma pasar de su establecimiento a las recomendaciones para el manejo, el desarrollo de nuevos enfoques como los análisis moleculares y herramientas como la PG se han comenzado a ver como una aproximación valiosa para reducir los ciclos generacionales y hacer predicciones en entornos no estudiados (Ratcliffe et al., 2019; Thistlethwaite et al., 2019). De esta manera, se busca implementar planes de manejo y conservación que mitiguen los efectos del CC y ayuden a conservar la diversidad genética enfocándose en la resiliencia de los ecosistemas (Aitken et al., 2008; Aitken y Whitlock, 2013). Sin embargo, si bien son útiles y complementarios, los enfoques de PG para preseleccionar individuos para reforestar y recuperar los ecosistemas degradados no pueden sustituir los ensayos de procedencias, que aún son necesarios para estudiar las respuestas plásticas y adaptativas de los árboles al cambio climático (Hansen et al., 2012; Koskela et al., 2014; Pluess et al., 2016).

1.8 Sitio de estudio

La presente tesis se llevó a cabo en el Área de Protección de Flora y Fauna Nevado de Toluca (APFFNT), que es una de las áreas de protección más importantes de México. Cuenta con una

gran riqueza natural y una estructura orográfica particular, siendo que allí se forma la cuenca hidrográfica del Lerma, una de las más importantes del país por su provisión de agua dulce para las comunidades humanas de la región. Además, presenta extensos bosques templados que contribuyen a la captura y almacén de carbono, así como al origen y protección de los recursos del suelo (Arzate-Fernández et al., 2015).

El APFFNT está ubicada en el Estado de México (Fig. 1), al suroeste del valle de Toluca, dentro de la provincia fisiográfica Sistema Volcánico Transversal. Está localizada entre los 18°59' y 19°13' N y los 99°37' y 99°58' W. El volcán Nevado de Toluca representa la cuarta montaña más alta del país, alcanzando su altura máxima a los 4,660 m.s.n.m. Presenta un intervalo altitudinal que incluye al cono volcánico y otra serie de geoformas que se extienden hacia el noroeste (Candeau y Franco, 2007; Domínguez-Tejeda et al., 2010). El APFFNT abarca aproximadamente 50 mil hectáreas, incluyendo parte de los municipios de Toluca, Zinacantepec, Amanalco de Becerra, Almoloya de Juárez, Temascaltepec, Coatepec Harinas, Villa Guerrero, Calimaya y Tenango del Valle. Se caracteriza por un clima templado frío con una temperatura media anual de 13.1°C y una precipitación media anual de 1,219. mm. Los suelos son derivados de afloramientos andesíticos, producto del intemperismo de cenizas volcánicas con una edad aproximada de 1.6 a 1.2 millones de años (Arzate-Fernández et al., 2015).

En las últimas décadas esta APFFNT se ha venido deteriorando por las actividades antrópicas (principalmente ganadería, agricultura, extracción selectiva de madera y tala ilegal) debido a su cercanía con áreas urbanizadas (CONANP, 2007; Granados-Ramírez et al., 2017). Dada su categoría de APFF, el Consejo Civil Mexicano para la Silvicultura Sostenible (CCMSS) ofrece a los habitantes locales opciones para cultivar y manejar el bosque; siendo uno de los objetivos principales el repoblar los sitios (o rodales forestales) mediante introducción de individuos por reforestación con especies nativas y brindar alternativas productivas viables para los pobladores. Esto también permite reducir el impacto humano y mitigación del CC promoviendo la conservación de los ecosistemas boscosos (CONANP, 2013). A pesar de las constantes presiones para el ecosistema, el paisaje manejado antropicamente nos permite realizar estudios genéticos y evolutivos, mismos que nos pueden proporcionar información valiosa para proponer planes de manejo y preservación de la diversidad genética (Candeau y Franco, 2007; Sáenz-Romero et al., 2003).

La diversidad de recursos naturales es, sin duda, una característica particular de esta región, los bosques templados son los que predominan en esta zona y estructuran las poblaciones primarias de este ecosistema, integradas por tres tipos de comunidades: el bosque de oyamel (*Abies religiosa*), el bosque de pino (conformado por tres especies, *Pinus montezumae*, *Pinus hartwegii* y *Pinus pseudostrobus*) y el bosque mixto (formado por pino, oyamel y encino (*Quercus* spp; Arzate-Fernández et al., 2015).

Fig. 1 Mapa de la ubicación y área del Nevado de Toluca en el Estado de México, México. En escala de azules se presentan algunos de los rodales forestales trabajados por el CCMSS. Los puntos indican la ubicación de los dos rodales estudiados en esta tesis: Rincón de Guadalupe (rojo) y San Bartolo (naranja).

1.9 Especie de estudio - *Abies religiosa*

Abies religiosa (Kunth) Schltdl. & Cham. (oyamel) es una conífera (Pinaceae) ampliamente distribuida en la zona ecológica templada subhúmeda del centro de México (entre 2800 y 3500 m.s.n.m), principalmente en el estrato superior del Sistema Volcánico Transversal (Rzedowski y Fryxell, 2006), a lo largo de un gradiente longitudinal que va de -96° a -104°. La especie puede crecer y desarrollarse principalmente en suelos jóvenes y profundos derivados de ceniza volcánica, con abundante MO y C y con alta humedad (Méndez-González et al., 2017). Los bosques de oyamel forman poblaciones densas y generalmente monoespecíficas, distribuidas en parches relictuales y dispersos entre sí (Rzendowski, 1978). Estos provocan un ambiente de sombra y humedad que permite la proliferación de hongos, musgos, helechos y hepáticas (CONANP, 2013). Sus poblaciones se han diferenciado históricamente en respuesta a la variación ambiental (Cruz-Nicolás et al., 2020), aparentemente siguiendo las características climáticas locales (p. ej. precipitación anual) y respondiendo a un gradiente de elevación que se traduce en variación fenotípica en diferentes rasgos (tamaño del cono y longitud de las acículas; Ortiz-Bibian et al., 2017). Se ha sugerido que esta diferenciación tiene una base genética adaptativa en respuesta al ambiente (Sáenz-Romero et al., 2016; Ortiz-Bibian et al., 2017). Es posible, además, que la heterogeneidad en la composición edáfica y las propiedades físicas del suelo (p. ej., la cantidad de materia orgánica) también hayan conllevado a una adaptación microambiental (Méndez-González et al., 2017), lo que hace suponer una interacción G × E. De esta forma, la cantidad de variación genética reportada podría considerarse como una estrategia adaptativa para colonizar suelos heterogéneos (o simplemente otros efectos como plasticidad fenotípica).

Con los aumentos de calor y sequía predichos para las próximas décadas, asociados al CC, se espera que el nicho adecuado de *A. religiosa* en el centro de México se reduzca entre el 69 y 97% de aquí al año 2090; lo que hará susceptibles a las poblaciones al estrés por desecación y a la incidencia de plagas, como los insectos descortezadores y las plantas parásitas (Sáenz-Romero et al., 2012). Por lo tanto, esta variación ambiental obliga a los investigadores y mejoradores genéticos a tomar medidas preventivas en el manejo de estos bosques que permitan implementar enfoques de adaptación al cambio climático(Ortiz-Bibian et al., 2017).

Se ha propuesto que la reforestación con individuos de oyamel en un ambiente propicio para su crecimiento, desarrollo y reproducción, puede representar una oportunidad para mitigar los efectos adversos del CC y conservar su ecosistema natural (Sáenz-Romero et al., 2016). Esta

estrategia de manejo consiste en transferir germoplasma dentro de poblaciones intervenidas para reactivar con mayor eficiencia los servicios ecosistémicos, como la recuperación de suelos degradados y conservar la biodiversidad (Carbajal-Navarro et al., 2019; Cruzado-Vargas et al., 2019).

Abies religiosa es el centro de estudio de esta tesis debido a que exhibe características importantes para evaluar modelos poligénicos y cuantificar su variación fenotípica y los procesos adaptativos subyacentes. En particular, se espera que el origen de los árboles para reforestación sea un factor importante para detectar los efectos de las diferencias genéticas sobre la variación fenotípica. A diferencia de las especies con ciclos de vida corto, y teniendo en cuenta que las reforestaciones son de regiones cercanas (*com. pers*. Andrés Juárez), se espera que las plantas locales y las utilizadas en reforestación tengan una diversidad genética semejante, ya que el manejo forestal en el Nevado de Toluca es reciente. En comparación con las plantas de regeneración natural, se espera que las plantas reforestadas tengan un crecimiento más lento (al menos en los primeros años) (Ledig y Kitzmiller, 1992; Koskela et al., 2014), sugiriendo que en las primeras etapas de crecimiento, los procesos adaptativos pueden operar de diferente forma entre grupos. Esto permitirá evaluar la estructura genética y su asociación con la variación fenotípica, posterior a la reforestación. Sin embargo, también es importante considerar los efectos de las presiones selectivas de la heterogeneidad del suelo y la respuesta de interacción G × E.

Al ser una especie forestal amenazada por la degradación de sus ecosistemas y el CC, se puede suponer que los efectos de la variación genética y la heterogenidad del suelo influenciarán en la variación del fenotipo. Estos efectos abren las puertas a la exploración e implementación de modelos poligénicos en poblaciones naturales. Ya que los bosques de oyamel se caracterizan por presentar suelos heterogéneos en lugares con condiciones ambientales similares (Méndez-González et al., 2017), se espera detectar cambios en los rasgos cuantitativos y en la respuesta adaptativa entre plantas locales e introducidas. Además, se espera poder modelar y predecir la relación entre la diversidad genética existente y la variación fenotípica vinculada a ésta, todo a partir de modelos estadísticos de PG para promover perspectivas de manejo y conservación.

1.10 Referencias

Aitken, S. N., and Bemmels, J. B. (2016). Time to get moving: Assisted gene flow of forest trees. *Evolutionary Applications* 9, 271–290. doi:10.1111/eva.12293.

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, 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, 95–111. doi:10.1111/j.1752-4571.2007.00013.x.
Alberto, F. J., Aitken, S. N., Alía, R., González-Martínez, S. C., Hänninen, H., Kremer, A., et al. (2013). Potential for evolutionary responses to climate change - evidence from tree populations. *Global Change Biology* 19, 1645–1661. doi:10.1111/gcb.12181.
Alimi, N. A. (2016). *Statistical methods for QTL mapping and genomic prediction of multiple traits and environments: case studies in pepper*. Available at: http://edepot.wur.nl/390205.
Badenes, M. L., Fernández i Martí, A., Ríos, G., and Rubio-Cabetas, M. J. (2016). Application of genomic technologies to the breeding of trees. *Frontiers in Genetics* 7, 1–13. doi:10.3389/fgene.2016.00198.
Bailey, J. K., Schweitzer, J. A., Úbeda, F., Koricheva, J., LeRoy, C. J., Madritch, M. D., et al. (2009). From genes to ecosystems: A synthesis of the effects of plant genetic factors across levels of organization. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1607–1616. doi:10.1098/rstb.2008.0336.
Ballesta, P., Bush, D., Silva, F. F., and Mora, F. (2020). Genomic predictions using low-density SNP markers, pedigree and GWAS information: A case study with the non-model species Eucalyptus cladocalyx. *Plants* 9. doi:10.3390/plants9010099.
Ballesta, P., Serra, N., Guerra, F. P., Hasbún, R., and Mora, F. (2018). Genomic prediction of growth and stem quality traits in Eucalyptus globulus Labill. at its southernmost distribution limit in Chile. *Forests* 9, 1–18. doi:10.3390/f9120779.
Bartholomé, J., Van Heerwaarden, J., Isik, F., Boury, C., Vidal, M., Plomion, C., et al. (2016). Performance of genomic prediction within and across generations in maritime pine. *BMC Genomics* 17, 1–14. doi:10.1186/s12864-016-2879-8.
Beaulieu, J., Doerksen, T., Clément, S., Mackay, J., and Bousquet, J. (2014a). Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. *Heredity* 113, 343–352. doi:10.1038/hdy.2014.36.
Beaulieu, J., Doerksen, T. K., Mackay, J., Rainville, A., and Bousquet, J. (2014b). Genomic selection accuracies within and between environments and small breeding groups in white spruce. 1–16.
Black, W. C., Baer, C. F., Antolin, M. F., and Duteau, N. M. (2001). POPULATION G ENOMICS : Genome-Wide.
Blanquart, F., Kaltz, O., Nuismer, S. L., and Gandon, S. (2013). A practical guide to measuring local adaptation. *Ecology Letters* 16, 1195–1205. doi:10.1111/ele.12150.
Breshears, D. D., Adams, H. D., Eamus, D., Mcdowell, N. G., Law, D. J., Will, R. E., et al. (2013). The critical amplifying role of increasing atmospheric moisture demand on tree mortality and associated regional die-off. *Frontiers in Plant Science* 4, 2–5. doi:10.3389/fpls.2013.00266.
Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S. M., et al. (2018). Global trait–environment relationships of plant communities. *Nature Ecology and Evolution* 2, 1906–1917. doi:10.1038/s41559-018-0699-8.
Candeau, D. R. y Franco M. S. (2007). Dinámica y condiciones de vida de la población del parque nacional nevado de toluca (PNNT) en la generación de presión a los ecosistemas circundantes y de impactos ambientales a través de un sistema de información geográ ca. Investigaciones geográ cas, (62), 44-68.
Cappai, C., Kemanian, A. R., Lagomarsino, A., Roggero, P. P., Lai, R., Agnelli, A. E., et al. (2017). Small-scale spatial variation of soil organic matter pools generated by cork oak trees in Mediterranean agro-silvo-pastoral systems. *Geoderma* 304, 59–67. doi:10.1016/j.geoderma.2016.07.021.
Carbajal-Navarro, A., Navarro-Miranda, E., Blanco-García, A., Cruzado-Vargas, A. L., Gómez-Pineda, E., Zamora-Sánchez, C., et al. (2019). Ecological Restoration of Abies religiosa Forests Using Nurse Plants and Assisted Migration in the Monarch Butterfly Biosphere Reserve, Mexico. *Frontiers in Ecology and Evolution* 7. doi:10.3389/fevo.2019.00421.
Carvalho, C. S., Forester, B. R., Mitre, S. K., Alves, R., Imperatriz-Fonseca, V. L., Ramos, S. J., et al. (2020). *Combining genotype, phenotype, and environmental data to delineate site-adjusted provenance strategies for ecological restoration*. doi:10.1111/1755-0998.13191.
Charmantier A., G. D. y K. L. (2014). Quantitative Genetics in the wild. *Current Genomics*. doi:10.2174/138920212800543110.
Chen, J., Shen, W., Xu, H., Li, Y., and Luo, T. (2019). The composition of nitrogen-fixing microorganisms correlates with soil nitrogen content during reforestation: A comparison between legume and non-legume plantations. *Frontiers in Microbiology* 10, 1–11. doi:10.3389/fmicb.2019.00508.
Chen, Z. Q., Baison, J., Pan, J., Karlsson, B., Andersson, B., Westin, J., et al. (2018). Accuracy of genomic selection

for growth and wood quality traits in two control-pollinated progeny trials using exome capture as the genotyping platform in Norway spruce. *BMC Genomics* 19. doi:10.1186/s12864-018-5256-y.

Collevatti, R. G., Novaes, E., Silva-Junior, O. B., Vieira, L. D., Lima-Ribeiro, M. S., and Grattapaglia, D. (2019). A genome-wide scan shows evidence for local adaptation in a widespread keystone Neotropical forest tree. *Heredity* 123, 117–137. doi:10.1038/s41437-019-0188-0.

Comisión Nacional de Áreas Naturales Protegidas (en línea). www.conanp.gob.mx Julio 3, 2007.

Comisión Nacional de Áreas Naturales Protegidas (en línea). www.conanp.gob.mx noviembre 15, 2013.

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., et al. (2017). Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends in Plant Science* 22, 961–975. doi:10.1016/j.tplants.2017.08.011.

Cruzado-Vargas, A. L., Zamudio-Sánchez, F. J., Rodríguez-Yam, G. A., Carbajal-Navarro, A. L., Blanco-García, J. A., and Sáenz-Romero, C. (2019). Growth of naturally regenerated Abies religiosa (Kunth) Schltdl. & Cham. Seedlings in a nursery and genetic variation among provenances. *Revista Chapingo, Serie Ciencias Forestales y del Ambiente* 26, 85–96. doi:10.5154/r.rchscfa.2019.01.013.

Daniels, J. D. (1984). Role of tree improvement in intensive forest management. *Forest Ecology and Management* 8, 161–195. doi:10.1016/0378-1127(84)90052-5.

de los Campos, G., Sorensen, D., and Gianola, D. (2015). Genomic Heritability: What Is It? *PLoS Genetics* 11, 1–21. doi:10.1371/journal.pgen.1005048.

De Mita, S., Thuillet, A. C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., et al. (2013). Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology* 22, 1383–1399. doi:10.1111/mec.12182.

de Villemereuil, P., Schielzeth, H., Nakagawa, S., and Morrissey, M. (2016). General methods for evolutionary quantitative genetic inference from generalized mixed models. *Genetics* 204, 1281–1294. doi:10.1534/genetics.115.186536.

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, 5–29. doi:10.1146/annurev-ecolsys-110512-135806.

Domínguez-Tejeda, E. M., Flores, O. P., Muciño, A. M. e. y Sierra-Domínguez, G. (2010). El Parque Nacional Nevado de Toluca y la falta de investigación para la toma de decisiones, caso específico: fauna. Ciencias Agrícolas Informa, 19, 42-52.

Elliot, W. J., Page-Dumroese, D., and Robichaud, P. R. (2019). The Effects of Forest Management on Erosion and Soil Productivity*. *Soil Quality and Soil Erosion*, 195–208. doi:10.1201/9780203739266-12.

Ellison, D., Morris, C. E., Locatelli, B., Sheil, D., Cohen, J., Murdiyarso, D., et al. (2017). Trees, forests and water: Cool insights for a hot world. *Global Environmental Change* 43, 51–61. doi:10.1016/j.gloenvcha.2017.01.002.

FAO (2014). *World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps*. doi:10.1017/S0014479706394902.

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, 2215–2233. doi:10.1111/mec.14584.

George, J. P., Grabner, M., Karanitsch-Ackerl, S., Mayer, K., Weißenbacher, L., and Schueler, S. (2017). Genetic variation, phenotypic stability, and repeatability of drought response in European larch throughout 50 years in a common garden experiment. *Tree Physiology* 37, 33–46. doi:10.1093/treephys/tpw085.

Goddard, M. E., and Hayes, B. J. (2009). Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature Publishing Group* 10, 381–391. doi:10.1038/nrg2575.

González-Camacho, J. M., Ornella, L., Pérez-Rodríguez, P., Gianola, D., Dreisigacker, S., and Crossa, J. (2018). Applications of machine learning methods to genomic selection in breeding wheat for rust resistance. *Plant Genome* 11, 1–15. doi:10.3835/plantgenome2017.11.0104.

Grattapaglia, D. (2014). *Genomics of plant genetic resources: Volume 1. Managing, sequencing and mining genetic resources. Chapter 26 Breeding Forest Trees by Genomic Selection: Current Progress and the Way Forward*. doi:10.1007/978-94-007-7572-5.

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

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

Grattapaglia, D., Silva-Junior, O. B., Resende, R. T., Cappa, E. P., Müller, B. S. F., Tan, B., et al. (2018). Quantitative genetics and genomics converge to accelerate forest tree breeding. *Frontiers in Plant Science* 871, 1–10.

doi:10.3389/fpls.2018.01693.
Guerrero, J., Andrello, M., Burgarella, C., and Manel, S. (2018). Soil environment is a key driver of adaptation in Medicago truncatula: new insights from landscape genomics. *New Phytologist* 219, 378–390. doi:10.1111/nph.15171.
Hansen, J., Sato, M., and Ruedy, R. (2012). Perception of climate change. *Proceedings of the National Academy of Sciences of the United States of America* 109. doi:10.1073/pnas.1205276109.
Hayes, B. J., Bowman, P. J., Chamberlain, A. J., and Goddard, M. E. (2009). Invited review : Genomic selection in dairy cattle : Progress and challenges. *Journal of Dairy Science* 92, 433–443. doi:10.3168/jds.2008-1646.
Hodge, G. R., and Dvorak, W. S. (2015). Provenance variation and within-provenance genetic parameters in Eucalyptus urophylla across 125 test sites in Brazil, Colombia, Mexico, South Africa and Venezuela. *Tree Genetics and Genomes* 11. doi:10.1007/s11295-015-0889-3.
Holliday, J. A., Aitken, S. N., Cooke, J. E. K., Fady, B., Gonz Alez-Martinez, S. C., Heuertz, M., et al. (2017). Advances in ecological genomics in forest trees and applications to genetic resources conservation and breeding. *Molecular Ecology* 26, 706–717. doi:10.1111/mec.13963.
Hooper, D. U., Adair, E. C., Cardinale, B. J., Byrnes, J. E. K., Hungate, B. A., Matulich, K. L., et al. (2012). A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature* 486, 105–108. doi:10.1038/nature11118.
Housset, J. M., Nadeau, S., Isabel, N., Depardieu, C., Duchesne, I., Lenz, P., et al. (2018). Tree rings provide a new class of phenotypes for genetic associations that foster insights into adaptation of conifers to climate change. *New Phytologist* 218, 630–645. doi:10.1111/nph.14968.
Ikram, S., and Chardon, F. (2010). Plant Quantitative Traits. *Encyclopedia of Life Sciences*. doi:10.1002/9780470015902.a0002021.pub2.
Isabel, N., Holliday, J. A., and Aitken, S. N. (2020). Forest genomics: Advancing climate adaptation, forest health, productivity, and conservation. *Evolutionary Applications* 13, 3–10. doi:10.1111/eva.12902.
Isik, F. (2014). Genomic selection in forest tree breeding: The concept and an outlook to the future. *New Forests* 45, 379–401. doi:10.1007/s11056-014-9422-z.
Isik, F., Bartholomé, J., Farjat, A., Chancerel, E., Raffin, A., Sanchez, L., et al. (2016). Plant Science Genomic selection in maritime pine. *Plant Science* 242, 108–119. doi:10.1016/j.plantsci.2015.08.006.
Iwata, H., Minamikawa, M. F., Kajiya-Kanegae, H., Ishimori, M., and Hayashi, T. (2016). Genomics-assisted breeding in fruit trees. *Breeding Science* 66, 100–115. doi:10.1270/jsbbs.66.100.
Kahlke, R. M., and Hon, B. A. (2014). International Journal of Qualitative Methods-2014-Kahlke-37-52. *International Journal of Qualitative Methods* 13, 37–52.
Koskela, J., Vinceti, B., Dvorak, W., Bush, D., Dawson, I. K., Loo, J., et al. (2014). Utilization and transfer of forest genetic resources: A global review. *Forest Ecology and Management* 333, 22–34. doi:10.1016/j.foreco.2014.07.017.
Kubota, S., Iwasaki, T., Hanada, K., Nagano, A. J., Fujiyama, A., Toyoda, A., et al. (2015). A Genome Scan for Genes Underlying Microgeographic-Scale Local Adaptation in a Wild Arabidopsis Species. *PLoS Genetics* 11, 1–26. doi:10.1371/journal.pgen.1005361.
Lafleur, B., Labrecque, M., Arnold, A. A., and Bélanger, N. (2015). Organic carbon accumulation in topsoil following afforestation with willow: Emphasis on leaf litter decomposition and soil organic matter quality. *Forests* 6, 769–793. doi:10.3390/f6030769.
Lascoux, M., Glémin, S., and Savolainen, O. (2016). Local Adaptation in Plants. *eLS*, 1–7. doi:10.1002/9780470015902.a0025270.
Le, H. D., Smith, C., Herbohn, J., and Harrison, S. (2012). More than just trees: Assessing reforestation success in tropical developing countries. *Journal of Rural Studies* 28, 5–19. doi:10.1016/j.jrurstud.2011.07.006.
Lenz, P. R. N., Beaulieu, J., Mansfield, S. D., Clément, S., Desponts, M., and Bousquet, J. (2017). Factors affecting the accuracy of genomic selection for growth and wood quality traits in an advanced-breeding population of black spruce (Picea mariana). *BMC Genomics* 18, 1–17. doi:10.1186/s12864-017-3715-5.
Lenz, P. R. N., Nadeau, S., Mottet, M. J., Perron, M., Isabel, N., Beaulieu, J., et al. (2020). Multi-trait genomic selection for weevil resistance, growth, and wood quality in Norway spruce. *Evolutionary Applications* 13, 76–94. doi:10.1111/eva.12823.
Li, Y., Yang, K., Yang, W., Chu, L., Chen, C., Zhao, B., et al. (2017). Identification of QTL and qualitative trait loci for agronomic traits using SNP markers in the Adzuki bean. *Frontiers in Plant Science* 8. doi:10.3389/fpls.2017.00840.
Lin, Y. Te, Whitman, W. B., Coleman, D. C., and Chiu, C. Y. (2018). Effects of reforestation on the structure and diversity of bacterial communities in subtropical low Mountain Forest Soils. *Frontiers in Microbiology* 9, 1–

10. doi:10.3389/fmicb.2018.01968.
Lin, Z., Hayes, B. J., and Daetwyler, H. D. (2014). Genomic selection in crops, trees and forages: A review. *Crop and Pasture Science* 65, 1177–1191. doi:10.1071/CP13363.
Liu, M., Sui, X., Hu, Y., and Feng, F. (2019). Microbial community structure and the relationship with soil carbon and nitrogen in an original Korean pine forest of Changbai Mountain, China. *BMC microbiology* 19, 218. doi:10.1186/s12866-019-1584-6.
Lobell, D. B., and Gourdji, S. M. (2012). The influence of climate change on global crop productivity. *Plant Physiology* 160, 1686–1697. doi:10.1104/pp.112.208298.
Madritch, M. D., Greene, S. L., and Lindroth, R. L. (2009). Genetic mosaics of ecosystem functioning across aspen-dominated landscapes. *Oecologia* 160, 119–127. doi:10.1007/s00442-009-1283-3.
Madritch, M., Donaldson, J. R., and Lindroth, R. L. (2006). Genetic identity of Populus tremuloides litter influences decomposition and nutrient release in a mixed forest stand. *Ecosystems* 9, 528–537. doi:10.1007/s10021-006-0008-2.
Martins, K., Gugger, P. F., Llanderal-Mendoza, J., González-Rodríguez, A., Fitz-Gibbon, S. T., Zhao, J. L., et al. (2018). Landscape genomics provides evidence of climate-associated genetic variation in Mexican populations of Quercus rugosa. *Evolutionary Applications* 11, 1842–1858. doi:10.1111/eva.12684.
Méndez-González, I. D., Jardón-Barbolla, L., and Jaramillo-Correa, J. P. (2017). Differential landscape effects on the fine-scale genetic structure of populations of a montane conifer from central Mexico. *Tree Genetics and Genomes* 13. doi:10.1007/s11295-017-1112-5.
Merilä, J., and Hendry, A. P. (2014). Climate change, adaptation, and phenotypic plasticity: The problem and the evidence. *Evolutionary Applications* 7, 1–14. doi:10.1111/eva.12137.
Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps.
Meuwissen, T., Hayes, B., and Goddard, M. (2016). Genomic selection: A paradigm shift in animal breeding. *Animal Frontiers* 6, 6–14. doi:10.2527/af.2016-0002.
Moran, E., Lauder, J., Musser, C., Stathos, A., and Shu, M. (2017). The genetics of drought tolerance in conifers. *New Phytologist* 216, 1034–1048. doi:10.1111/nph.14774.
Müller, B. S. F., Neves, L. G., de Almeida Filho, J. E., Resende, M. F. R., Muñoz, P. R., dos Santos, P. E. T., et al. (2017). Genomic prediction in contrast to a genome-wide association study in explaining heritable variation of complex growth traits in breeding populations of Eucalyptus. *BMC Genomics* 18, 1–17. doi:10.1186/s12864-017-3920-2.
Munoz, P. R., Resende, M. F. R., Huber, D. A., Quesada, T., Resende, M. D. V., Neale, D. B., et al. (2014). Genomic relationship matrix for correcting pedigree errors in breeding populations: Impact on genetic parameters and genomic selection accuracy. *Crop Science* 54, 1115–1123. doi:10.2135/cropsci2012.12.0673.
Muthoo, M. (2002). Mountain environment and development. *Unasylva* 53, 26–35.
Nacry, P., Bouguyon, E., and Gojon, A. (2013). Nitrogen acquisition by roots: Physiological and developmental mechanisms ensuring plant adaptation to a fluctuating resource. *Plant and Soil* 370, 1–29. doi:10.1007/s11104-013-1645-9.
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, 8649–8664. doi:10.1002/ece3.2550.
Nakaya, A., and Isobe, S. N. (2012). Will genomic selection be a practical method for plant breeding? *Annals of Botany* 110, 1303–1316. doi:10.1093/aob/mcs109.
Neale, D. B., and Kremer, A. (2011). Forest tree genomics: Growing resources and applications. *Nature Reviews Genetics* 12, 111–122. doi:10.1038/nrg2931.
Oddou-Muratorio, S., Davi, H., and Lefèvre, F. (2020). Integrating evolutionary, demographic and ecophysiological processes to predict the adaptive dynamics of forest tree populations under global change. *Tree Genetics and Genomes* 16, 1–22. doi:10.1007/s11295-020-01451-1.
Ortiz-Bibian, M. A., Blanco-García, A., Lindig-Cisneros, R. A., Gómez-Romero, M., Castellanos-Acuña, D., Herrerías-Diego, Y., et al. (2017). Genetic Variation in Abies religiosa for Quantitative Traits and Delineation of Elevational and Climatic Zoning for Maintaining Monarch Butterfly Overwintering Sites in Mexico, considering Climatic Change. *Silvae Genetica* 66, 14–23. doi:10.1515/sg-2017-0003.
Plomion, C., Bastien, C., Bogeat-Triboulot, M. B., Bouffier, L., Déjardin, A., Duplessis, S., et al. (2016). Forest tree genomics: 10 achievements from the past 10 years and future prospects. *Annals of Forest Science* 73, 77–103. doi:10.1007/s13595-015-0488-3.
Pluess, A. R., Frank, A., Heiri, C., Lalagüe, H., Vendramin, G. G., and Oddou-Muratorio, S. (2016). Genome-

environment association study suggests local adaptation to climate at the regional scale in Fagus sylvatica. *New Phytologist* 210, 589–601. doi:10.1111/nph.13809.
Potter, R. S., and Snyder, R. S. (1915). The determination of ammonia in soils. *Industrial and Engineering Chemistry* 7, 221–226. doi:10.1021/ie50075a021.
Pregitzer, C. C., Bailey, J. K., and Schweitzer, J. A. (2013). Genetic by environment interactions affect plant-soil linkages. *Ecology and Evolution* 3, 2322–2333. doi:10.1002/ece3.618.
Pugnaire, F. I., Morillo, J. A., Peñuelas, J., Reich, P. B., Bardgett, R. D., Gaxiola, A., et al. (2019). Climate change effects on plant-soil feedbacks and consequences for biodiversity and functioning of terrestrial ecosystems. *Science Advances* 5, eaaz1834. doi:10.1126/sciadv.aaz1834.
Rae, J. G. (2013). Abiotic factors affect microhabitat selection and community dynamics in a sandy-bottom lotic chironomid midge assemblage. *Hydrobiologia* 700, 121–130. doi:10.1007/s10750-012-1223-9.
Ratcliffe, B., Thistlethwaite, F. R., El-Dien, O. G., Cappa, E., Porth, I., Klápště, J., et al. (2019). Inter- and Intra-Generation Genomic Predictions for Douglas-fir Growth in Unobserved Environments. *bioRxiv*, 540765. doi:10.1101/540765.
Raven, J. A., and Andrews, M. (2010). Evolution of tree nutrition. *Tree Physiology* 30, 1050–1071. doi:10.1093/treephys/tpq056.
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, 4348–4370. doi:10.1111/mec.13322.
Resende, M. D. V., Resende, M. F. R., Sansaloni, C. P., Petroli, C. D., Missiaggia, A. A., Aguiar, A. M., et al. (2012a). Genomic selection for growth and wood quality in Eucalyptus: Capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytologist* 194, 116–128. doi:10.1111/j.1469-8137.2011.04038.x.
Resende, M. D. V., Resende, M. F. R., Sansaloni, C. P., Petroli, C. D., Missiaggia, A. A., Aguiar, A. M., et al. (2012b). Genomic selection for growth and wood quality in Eucalyptus: Capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytologist* 194, 116–128. doi:10.1111/j.1469-8137.2011.04038.x.
Resende, M. F. R., Muñoz, P., Acosta, J. J., Peter, G. F., Davis, J. M., Grattapaglia, D., et al. (2012c). Accelerating the domestication of trees using genomic selection: Accuracy of prediction models across ages and environments. *New Phytologist* 193, 617–624. doi:10.1111/j.1469-8137.2011.03895.x.
Rzedowski, J.Fryxell, P. a. (2006). Vegetación de México (índice). *Taxon* 31, 793. doi:10.2307/1219727.
Rzendowski, J. (1978). Bosque de coníferas. *Vegetación de México*, 295–327. doi:10.1007/s13398-014-0173-7.2.
Sáenz-Romero, C., Lindig-Cisneros, R. A., Joyce, D. G., Beaulieu, J., Bradley, J. S. C., and Jaquish, B. C. (2016). Assisted migration of forest populations for adapting trees to climate change. *Revista Chapingo, Serie Ciencias Forestales y del Ambiente* 22, 303–323. doi:10.5154/r.rchscfa.2014.10.052.
Sáenz-Romero, C., Snively, A. E. y Lindig-Cisneros, R. (2003). Conservation and restoration of pine forest genetic resources in Mexico. Silvae genetica, 52(5-6), 233-236.
Sallam, A., Alqudah, A. M., Dawood, M. F. A., Baenziger, P. S., and Börner, A. (2019). Drought stress tolerance in wheat and barley: Advances in physiology, breeding and genetics research. *International Journal of Molecular Sciences* 20. doi:10.3390/ijms20133137.
Sauer, T. J., James, D. E., Cambardella, C. A., and Hernandez-Ramirez, G. (2012). Soil properties following reforestation or afforestation of marginal cropland. *Plant and Soil* 360, 375–390. doi:10.1007/s11104-012-1258-8.
Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics* 14, 807–820. doi:10.1038/nrg3522.
Savolainen, O., Pyhäjärvi, T., and Knürr, T. (2007). Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proceedings of the National Academy of Sciences of the United States of America* 38, 1752–1756. doi:10.1073/pnas.0906182107.
Sawitri, Tani, N., Na'iem, M., Widiyatno, Indrioko, S., Uchiyama, K., et al. (2020). Potential of Genome-Wide association studies and Genomic Selection to improve productivity and quality of commercial timber species in tropical rainforest, a case study of Shorea platyclados. *Forests* 11. doi:10.3390/f11020239.
Schneider, S., Azar, C., Baethgen, W., Hope, C., Moss, R., Leary, N., et al. (2011). Overview of Impacts, Adaptation and Vulnerability to Climate Change. *Impacts, adapation and vulnerability*, 75–103. doi:10.1016/j.gloenvcha.2005.11.007.
Schwartz, M. W., Hellmann, J. J., McLachlan, J. M., Sax, D. F., Borevitz, J. O., Brennan, J., et al. (2012). Managed Relocation: Integrating the Scientific, Regulatory, and Ethical Challenges. *BioScience* 62, 732–743.

doi:10.1525/bio.2012.62.8.6.
Schweitzer, J. A., Fischer, D. G., Rehill, B. J., Wooley, S. C., Woolbright, S. A., Lindroth, R. L., et al. (2011). Forest gene diversity is correlated with the composition and function of soil microbial communities. *Population Ecology* 53, 35–46. doi:10.1007/s10144-010-0252-3.
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, 69–75. doi:10.1007/s13595-015-0471-z.
Silfver, T., Paaso, U., Rasehorn, M., Rousi, M., and Mikola, J. (2015). Genotype × herbivore effect on leaf litter decomposition in betula pendula saplings: Ecological and evolutionary consequences and the role of secondary metabolites. *PLoS ONE* 10, 1–15. doi:10.1371/journal.pone.0116806.
Sniezko, R. A., and Koch, J. (2017). Breeding trees resistant to insects and diseases: putting theory into application. *Biological Invasions* 19, 3377–3400. doi:10.1007/s10530-017-1482-5.
Soltys-Kalina, D., Plich, J., Strzelczyk-Żyta, D., Śliwka, J., and Marczewski, W. (2016). The effect of drought stress on the leaf relative water content and tuber yield of a half-sib family of 'Katahdin'-derived potato cultivars. *Breeding Science* 66, 328–331. doi:10.1270/jsbbs.66.328.
Sork, V. L. (2018). Genomic Studies of Local Adaptation in Natural Plant Populations. *Journal of Heredity* 109, 3–15. doi:10.1093/jhered/esx091.
Suontama, M., Mckinley, R., Dungey, H., Telfer, E., Graham, N., Stovold, T., et al. (2019). Ef fi ciency of genomic prediction across two Eucalyptus nitens seed orchards with different selection histories. 370–379. doi:10.1038/s41437-018-0119-5.
Tamaki, I., Setsuko, S., and Tomaru, N. (2008). Genetic variation and differentiation in populations of a threatened tree, Magnolia stellata: Factors influencing the level of within-population genetic variation. *Heredity* 100, 415–423. doi:10.1038/sj.hdy.6801097.
Tan, B., Grattapaglia, D., Martins, G. S., Ferreira, K. Z., Sundberg, B., and Ingvarsson, P. K. (2017). Evaluating the accuracy of genomic prediction of growth and wood traits in two Eucalyptus species and their F1hybrids. *BMC Plant Biology* 17, 1–15. doi:10.1186/s12870-017-1059-6.
Thistlethwaite, F. R., El-Dien, O. G., Ratcliffe, B., Klápště, J., Porth, I., Chen, C., et al. (2020). Linkage disequilibrium vs. pedigree: Genomic selection prediction accuracy in conifer species. *PLoS ONE* 15, 1–14. doi:10.1371/journal.pone.0232201.
Thistlethwaite, F. R., Ratcliffe, B., Klápště, J., Porth, I., Chen, C., Stoehr, M. U., et al. (2017). Genomic prediction accuracies in space and time for height and wood density of Douglas-fir using exome capture as the genotyping platform. *BMC Genomics* 18, 1–16. doi:10.1186/s12864-017-4258-5.
Thistlethwaite, F. R., Ratcliffe, B., Klápště, J., Porth, I., Chen, C., Stoehr, M. U., et al. (2019). Genomic selection of juvenile height across a single-generational gap in Douglas-fir. *Heredity* 122, 848–863. doi:10.1038/s41437-018-0172-0.
Thomas Ledig, F., and Kitzmiller, J. H. (1992). Genetic strategies for reforestation in the face of global climate change. *Forest Ecology and Management* 50, 153–169. doi:10.1016/0378-1127(92)90321-Y.
Ukrainetz, N. K., Ritland, K., and Mansfield, S. D. (2008). Identification of quantitative trait loci for wood quality and growth across eight full-sib coastal Douglas-fir families. *Tree Genetics and Genomes* 4, 159–170. doi:10.1007/s11295-007-0097-x.
Valladares, F., Matesanz, S., Guilhaumon, F., Araújo, M. B., Balaguer, L., Benito-Garzón, M., et al. (2014). The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecology Letters* 17, 1351–1364. doi:10.1111/ele.12348.
Van Eenennaam, A. L., Weigel, K. A., Young, A. E., Cleveland, M. A., and Dekkers, J. C. M. (2014). Applied Animal Genomics : Results from the Field. *Annu. Rev. Anim. Biosci.* 2, 105–139. doi:10.1146/annurev-animal-022513-114119.
Yousefi, S., Abbassi-Daloii, T., Kraaijenbrink, T., Vermaat, M., Mei, H., van 't Hof, P., et al. (2018). A SNP panel for identification of DNA and RNA specimens. *BMC Genomics* 19, 1–12. doi:10.1186/s12864-018-4482-7.
Zapata-Valenzuela, J., Whetten, R. W., Neale, D., McKeand, S., and Isik, F. (2013). Genomic estimated breeding values using genomic relationship matrices in a cloned population of loblolly pine. *G3: Genes, Genomes, Genetics* 3, 909–916. doi:10.1534/g3.113.005975.
Zhang, M., Suren, H., and Holliday, J. A. (2019). Phenotypic and Genomic Local Adaptation across Latitude and Altitude in Populus trichocarpa. *Genome biology and evolution* 11, 2256–2272. doi:10.1093/gbe/evz151.

OBJETIVOS DE INVESTIGACIÓN

Los objetivos de investigación de esta tesis son:

• Desarrollar modelos poligénicos en árboles jóvenes de oyamel y evaluar las diferencias fenotípicas, genéticas y microambientales entre un grupo reforestado y uno naturalmente regenerado (Capítulo II).

• Identificar *loci* asociados a la variación en la composición edáfica dentro un rodal forestal, utilizando análisis de asociación genotipo ambiente (GEA) utilizando GBS (capítulo II).

• Desarrollar una revisión de literatura sobre las investigaciones en SG y PG, con el fin de discutir y mostrar las posibilidades de aplicar estas herramientas en especies de ambientes naturales y en especies con largos tiempos generacionales (capítulo III).

• Evaluar el potencial de la secuenciación mediante genotipificación como plataforma para hacer predicción genómica en poblaciones naturales de *Abies religiosa* (capítulo IV).

• Evaluar la precisión de estas predicciones a través de modelos predictivos de fenotipos entre ambientes naturales (capítulo IV).

CAPÍTULO II

Las interacciones genotipo-suelo dentro de la población impulsan la variación fenotípica en un bosque de oyamel en recuperación del centro de México

Within-population genotype – soil interactions drive phenotypic variation in a recovering fir forest from central Mexico

ABSTRACT

Within-population phenotypic variation can be used as a proxy for adaptive potential and individual performance in forest plantations and natural populations. Such variation can be viewed as the sum of genetic factors and their interaction with environmental variation (e.g., soil characteristics), which lead similar genotypes to produce contrasting phenotypes in different environments, and introduced individuals to perform differently from local trees in plantations and test trials. Predicting this variation is of great interest to foresters working with commercially important species. Such is the case of sacred fir (*Abies religiosa*), a species that is an important source of wood and resin for communities living above 2,500 m asl in central Mexico. We determined the contribution of both genetic and soil factors for predicting phenotypic performance of local naturally regenerated (NR) and introduced by reforestation (RF) seedlings in a sacred fir trial test performed by local communities in the Nevado de Toluca National Park. In spite of a low genetic differentiation between plant sources, NR seedlings outperformed RF plants in terms of height, diameter and water use. According to our models, a large part of these differences could be explained by the interaction of genetic, management and edaphic factors within the trial site, with local genotypes using more efficiently the available soil nitrogen than the introduced ones; thus indicating that planting could be changing plant microenvironment and have a measurable phenotypic effect. Association studies between genetic and edaphic variation further suggested that local adaptation might also be occurring at small within-population scales. All such traits, when correctly integrated, may result in fair genome-based phenotypic predictions for small secluded natural populations, which could be at the base of better reforestation practices, and conservation and assisted migration programs.

Key words: *Abies*; phenotypic variation; reforestation; local adaptation; forest regeneration; genotype – environment interaction (G × E); plant-soil interactions.

2.1 INTRODUCTION

Reforestation is essential for the restoration of natural ecosystems and ecosystem services; some of its goals include the reinstatement of deforested areas, recovering soil fertility and conserving biodiversity (Firn et al., 2007; Le et al., 2012; Loo et al., 2014). Traditional reforestation programs usually rely on seedlings from top-mother-trees or noteworthy provenances (Tolkamp et al., 1999; Wills et al., 2017). This often leads to transferring germplasm from one place to another; a practice that could also lead to introducing maladapted individuals in the target population (Lawson and Michler, 2014; Sebastian-Azcona et al., 2020). Understanding the effects of transferring germplasm can help forestry practices, and also aid forecasting species' responses to environmental changes (Sáenz-Romero et al., 2012; Aitken and Whitlock, 2013). Reforestation involves multiple life-history stages, from nursery seed germination, to seedling transplantation and survival in natural conditions (Koskela et al., 2014). Among the ecological factors that can affect seedling establishment in a new site, the soil-plant-atmosphere interactions seem the most relevant (Pregitzer et al., 2013; Carrasco-Carballido et al., 2019). It is noteworthy though that genetic-based adaptability is often overlooked in reforestation programs and modelling studies (Gray and Hamann, 2011; Lin et al., 2018; Isabel et al., 2020). However, the capacity of plants to adapt to a new environment, where they will face novel biotic and abiotic pressures, is crucial, because it directly translates into individual's performance and survival (Aitken et al., 2008; Vizcaíno-Palomar et al., 2014; Zhang et al., 2019).

The action of natural selection on introduced plants can result in rapid evolutionary change and local adaptation (Mitchell-Olds et al., 2007; Blanquart et al., 2013; Sork, 2018). Several studies have reported local adaptation at various geographic scales in forest trees (e.g., Savolainen et al., 2013; Sork et al., 2013), generally following environmental gradients (e.g. Savolainen et al., 2007; De Mita et al., 2013; Mahony et al., 2019). Thus, detecting candidate genes involved in local adaptation has been at the base of population genomics studies for several years, particularly for predicting the fate of individuals/species undergoing environmental changes (i.e. Aitken and Whitlock, 2013; Jaramillo-Correa et al., 2015; Lotterhos and Whitlock, 2015). However, most of the methods currently used only evaluate local associations between individual genes and environmental variables (Forester et al., 2016; Isabel et al., 2020), and overlook gene-gene (i.e. polygenic adaptation) or gene-environment interactions (i.e. G × E; Rellstab et al., 2015; Pluess et al., 2016; Forester et al., 2018), including conditonal neutrality. Traditionally, for accurately

exploring these factors, reciprocal transplants are necessary, which could be difficult to establish for some long-lived species, including forest trees (Lu et al., 2016). Common garden experiments are a powerful tool for studying local adaptation, where geo-referenced genetic materials (provenances) are set in a common environment to evaluate fitness-surrogate traits, like survival or growth (Blanquart et al., 2013; Lascoux et al., 2016; Wachowiak et al., 2018). However, unless experiments are replicated in different environments (i.e. multi-testing common gardens), detecting genotype-environment association is complicated.

Environmental heterogeneity can be substantial in tropical and mountain ecosystems (Muthoo, 2002; Dufour et al., 2006), even at fine-geographic scales (i.e. within the same forest patch; Sauer et al., 2012; Cappai et al., 2017; Méndez-González et al., 2017). Edaphic variation is probably one of the most heterogeneous factors that may affect plant performance and survival in these environments (due to differences in pH, organic horizon depth and associated soil biota), particularly in the tropics, where soils often have different ages and composition and may be heavily lixiviated from heavy rainfall (Cruz-Ruiz et al., 2012; Li et al., 2012; Augusto et al., 2017). Empirical studies in both natural populations and common gardens have indeed shown that growth and physiological traits may depend on soil nutrient processes (e.g., nitrogen mineralization), and soil microbial community composition (John et al., 2007; Raven and Andrews, 2010; Schweitzer et al., 2011). Other factors, like aboveground productivity, leaf litter decomposition, nutrient cycling and carbon (C) sequestration seem further correlated to intraspecific genetic diversity (Bailey et al., 2009; Ren et al., 2016), which could affect individual survival. This implies that plant and ecosystem functional traits may be ultimately affected by both genotypic traits and soil fertility, suggesting that G × E at small within-population scales should be a potential driver of local adaptation, and a main component of individuals' fitness in forest trees (Eckert et al., 2012; Talbot et al., 2017; Mahony et al., 2019; Brousseau et al., 2020).

Among the soil factors that may be playing key selective roles at various spatial scales in forest ecosystems, nitrogen (N), and phosphorus (P) availability for plant demand seem particularly noteworthy (Kubota et al., 2015; Liu et al., 2019; Du et al., 2020). Understanding the interaction of these components with genetic factors and how they affect tree phenotype and mortality rate (Pregitzer et al., 2013; Silfver et al., 2015; Bennett and Klironomos, 2019) is still a pending task; as this information is used to predict the adaptive potential of reforested individuals

into natural populations for conservation, restoration or management purposes (Baer 2016; Sork, 2018; Isabel et al., 2020; Carvalho et al., 2021).

Forest restoration seeks to improve ecosystem productivity and reduce degradation (Le et al., 2012; Koskela et al., 2014; Enache et al., 2016). However, some operations during reforestation (especially when scraping and extracting the topsoil) can cause severe disturbance and significantly change the physical (Cambi et al., 2017) and hydraulic properties of the soil (Abdi et al., 2017). The extent and severity of such disturbances depend on several factors, like the type of digging equipment (Lee et al., 2020), the use of machinery, the frequency of its use, and the texture and structure of the soil (Abdi et al., 2017). For instance, careless practicing can compact both the soil surface and the subsoil (Lee et al., 2020), diminishing the infiltration rate, the hydraulic conductivity and the aerial porosity, and increasing soil bulk density (Lee et al., 2020; Poltorak et al., 2018) with consequences to plant growth..

Abies religiosa (Kunth) Schltdl. and Cham. (sacred fir) is a major component of the highland tropical forests from central Mexico, along the Trans-Mexican Volcanic Belt (TMVB). It grows at elevations between 2800-3500 m where it forms large, and mostly monospecific stands (Méndez-González et al., 2017). Populations are found on relatively young soils, mainly andosolic, derived from volcanic ash, and rich in organic matter (Rzedowski, 2006). *Abies religiosa* is distributed in isolated patches on individual mountains, displaying a strong genetic differentiation at the landscape scale (i.e. G_{ST}= 0.20–0.51; for SSR markers, Cruz-Nicolás et al., 2020). At the phenotypic level, populations further exhibit heritable differentiation at growth traits (seedling height and dry biomass) along environmental gradients (Sáenz-Romero et al., 2012; Ortíz-Bibian et al., 2017). Although there is intrapopulation genetic structure, likely driven by soil differences (Méndez-González et al., 2017), little is known about putative within-population G × E interactions, and how they affect individuals' phenotype, especially at such fine-geographic scales.

In this study, we looked for within-population G × E interactions in a pilot study conducted in natural conditions within an *A. religiosa* stand with naturally regenerated and introduced individuals by reforestation. We compared newly generated genomic and microenvironmental soil data with phenotypic traits addressing the following questions: (*i*) does within-population phenotypic variation have a genetic component? (*ii*) Can phenotypic differences be at least partially explained by the soil characteristics? (*iii*) Does integrating G x E (genotype × soil

microenvironment) interaction increase the explained phenotypic variance in predictive models? And if so, (*iv*) does this soil-genotype interaction account for the phenotypic differences observed between local and introduced individuals? We expect that the G × E models here generated will improve future forest management practices and even help guiding assisted migration programs, along with results from previous studies (e.g., Ortiz-Bibian et al., 2017; Carbajal-Navarro et al., 2019; Cruzado-Vargas et al., 2019).

2.2 MATERIALS AND METHODS

2.2.1 Experimental design and site

The study was conducted in a managed forest area in the municipality of Amanalco de Becerra (Mexico), within the buffer zone of the Nevado de Toluca Flora and Fauna Protection Area (NTFFPA), central Mexico. This protected area is located about 80 km west of Mexico City, in the Trans-Mexican Volcanic Belt (TMVB); and encompasses the stratovolcano Nevado de Toluca (4,690 m asl). It harbors several forested areas dominated by pines, firs and oaks. Management focuses in the sustainable use of timber from different 'ejidos' (communally owned and managed lands) with the assistance of *Consejo Civil Mexicano para la Silvicultura Sostenible* (CCMSS). Our sampling focused on the "Rincón de Guadalupe" forest stands (19º 15' 19,33" to 19º 15' 29,84" N and -99º 57' 19,33" to -99º 57' 27,83" W). This site extends over 1.8 ha at elevations between 2,820 – 2,954 m. According to the World Reference Base for Soil classification, soils are andosols derived from ash and other volcanic ejections, and sedimentary rocks; characterized by the accumulation of stable organo-mineral complexes such as allophane (Krasilnikov et al., 2013; INEGI, 2015). This site encompasses highland tropical forests dominated by sacred fir (*Abies religiosa*). The climate is humid (Lang aridity index = 69; Lang, 1920), with a mean annual temperature of 15.5ºC and mean annual precipitation of 1,077 mm (CONAGUA, 2019). Measurements carried out in this study with a Digital Lux Meter Model LX-101; Taiwan, indicate the site has a low light intensity at noon, ranging from 292 to 487 μmol m^{-2} s^{-1}.

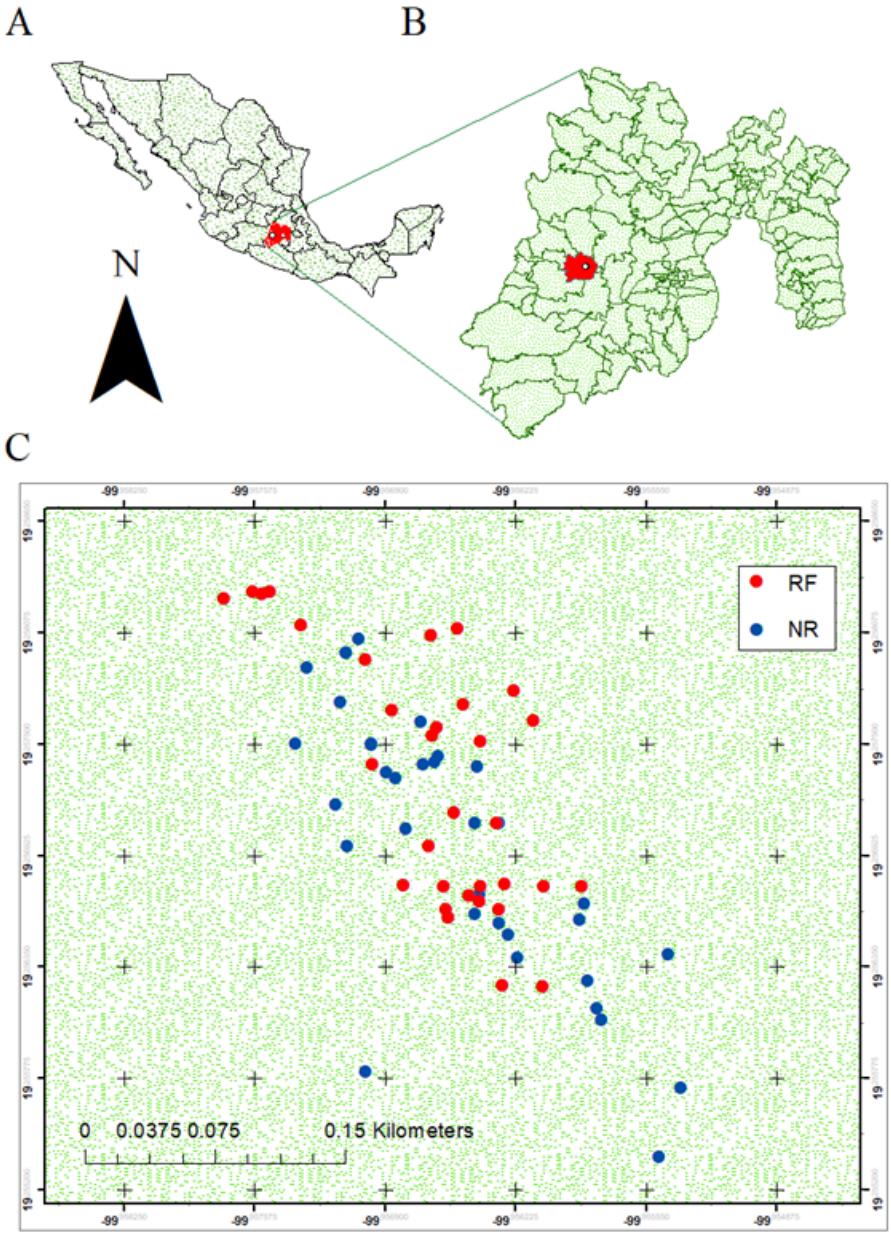

Fig. 1. Map of the study site showing. **A**. Location of the Mexican state where the investigation was performed (State of Mexico). **B.** location of the ejido 'Rincón de Guadalupe' within this state. **C.** Spatial distribution of sampled individuals. Each color represents an genetic origin class. RF, reforested introduced individuals; NR, naturally regenerated plants.

Selective wood extraction within the site has generated forest clearings that have been re-populated by natural regeneration (NR) from soil seed banks and introduced individuals through reforestation (RF). These individuals were planted in holes of approximately 26 cm in diameter and 30 cm deep. Soil was scraped, and the surface layer removed, along with most of the accumulated litter and organic matter. These plants are thus characterized by the absence of a

thick layer of organic matter around the stem. The planting hole was dug large enough to allow the plant to break through the soil and produce new roots. During reforestation, holes were slightly sunk as a supplementary water source for the planted trees. Although not exactly from the same site, reforested individuals originated from regional germplasm, as revealed in a previous population genomic study (Arenas et al. In press; see also below).

Sampling was performed in a 13 years-old stand, to ensure plants were about the same age (10.3 ± 0.48 years and 8.5 ± 0.15 years, mean ± SE for NR and RF, respectively,) and minimizing planting effects (but see below). The ages were obtained by counting the number of nodes per plant. We collected adult needles and soil, and measured phenotypic traits for 128 young trees between 1.2 and 2 m in height from both origins (51 NR- and 77 RF- trees) (Table S1); they grow intermixed along an east-to-west natural slope, under the same light and humidity conditions. All trees were devoid of apparent biotic damage and were at least 3 m apart from each other.

2.2.2. Analysis of phenotypic traits

2.2.2.1 Growth traits

We measured 12 morphological traits in the field, including total height (TH, cm), stem diameter at both the base (BD, cm) and 20 cm from the ground (SD, cm), crown radius (CR, cm) and length (CL, cm), average needle length (L_{leaf}, mm; estimated from 50 needles taken at random in 5 main branches at half of TH per individual; 250 in total), growth for the two previous years to the study (G_{16} and G_{17}, cm), and average growth for the previous four seasons (AG, cm). We inferred the age of each individual from bud-scar counts as in Hankin et al. (2018) and Urza and Sibold (2013). We then used general allometric equations to estimate the above-ground biomass in firs (AGB, kg) (1) (Chojnacky et al., 2014) and above-ground volume (AGV) (2) (Zianis et al., 2005).

$$ln\,(AGB) = \beta_0 + \beta_1 ln\,(SD)\ (1)$$

where β_0 = -2.3123 and β_1 =2.3482; and

$$AGV = \frac{\pi * CH * CR^2}{3} + \pi * SH * SR^2\ (2)$$

where CH is the crown height, CR the crown radius, SH is the height from the ground to where the branches start (or first branch) and SR the stem radius.

Although likely distorted because of a high proportion of compression wood during sampling (a distortion that should be uniform for all individuals sampled), specific wood density (WD,

g/cm^3) was determined for each plant by averaging measures from two pieces of wood collected from main branches. Wood density estimates from destructive sampling of main stems was not authorized by park authorities. We first measured water displacement within a graduated test tube to determine fresh volume of wood sample. We determined dry mass with an electronic balance (precision of 0.001g) after drying pieces in an oven at 75 ºC for 24 hours. WD was then estimated following Williamson and Wiemann (2010) as (3):

$$WD\ (\mathrm{g/cm^3}) = \frac{dry\ mass\ (\mathrm{g})}{fresh\ volume\ (\mathrm{cm^3})} \quad (3)$$

2.2.2.2 Physiological traits

Plants' performance (and phenotype), we further measured Four physiological variables were measured to assess plant performance (and phenotype). Vegetative tissue samples were taken between 9 am and 3 pm in the upper half of the crown of each individual, immediately placed on ice, and stored at -10 °C until further processing in the laboratory. Plant water status at the time of the study, water potential (Ψ, MPa) was first determined using the Schölander's pressure chamber technique (Boyer, 1967) and then, following Soltys-Kalina et al. (2016), we measured the leaf relative water content (RWC, %) as in (eq. 4). To do so, we used 20 randomly selected needles per plant, which were weighed by duplicate in an electronic balance (4).

$$RWC = \frac{fresh\ mass\ (\mathrm{g}) - dry\ mass\ (\mathrm{g})}{turgid\ mass\ (\mathrm{g}) - dry\ mass\ (\mathrm{g})} \times 100 \quad (4)$$

Specific leaf area (SLA, cm^2 g^{-1}) as a proxy of photosynthesis. SLA was defined in single-side as leaf area /dry mass (Poorter, 2002; Liu et al., 2017), and was estimated from the same needles above using images obtained with a scanner (HP Scanjet G3110). Leaf area of the images of each group of needles was calculated using *ImageJ* v.1.8 (Abràmoff et al., 2005).

Finally, we estimated the relative growth rate (RGR, $g\ g^{-1}d^{-1}$) as (5):

$$RGR = \frac{ln\ (AGB_{2017}) - ln\ (AGB_{2016})}{t_{2017} - t_{2016}} \quad (5)$$

Where AGB_{2016} and AGB_{2017} are the plant dry above-ground biomass for the two previous years to the study, and $t_{2017} - t_{2016}$ is the time spanned (in number of days) between the end of the previous growing season and the sampling date (Hoffmann and Poorter, 2002).

2.2.3 DNA sequencing, *de novo* assembly and genotyping

Total genomic DNA was extracted from young needles for a subset of 65 randomly selected plants (34 NR-individuals and 31 RF- individuals; **Fig. 1)**, following Telfer et al., (2013), and using a DNAeasy Plant Mini Kit (Qiagen, Germany). After DNA restriction with *PstI*, library preparation was carried out according to Poland et al. (2012). Single-end sequence reads were obtained from an Illumina HiSeq 2500 sequencer at the Institute of Integrative Biology and Systems at Université Laval (http://www.ibis.ulaval.ca/en/services-2/genomic-analysis-platform/). Demultiplexing and quality filtering, *de novo* assembly, read alignment and SNP calling were performed with an *Ipyrad* v0.7.23 (Eaton, 2014) pipeline. Assembling parameters included a clustering threshold of 0.9, a mindepth of 8, and a maximum barcode mismatch of 0 defining 80 bp reads after cutting them by quality. Both types of individuals (NR and RF) had to be represented at least once for a SNP to be called. Subsequently, monomorphic reads, variants with missing call rates above 25%, and samples with minimum allelic frequencies (MAF) below 5% and in Hardy-Weinberg disequilibrium ($P < 1x10^{-6}$) (Minamikawa et al., 2018) were eliminated with Plink v1.07 (Purcell, 2010).

2.2.3 Soil sampling and analyses

After removing the surface litter down to the mineral soil?, four topsoil (0-15 cm in depth) samples separated by 90° were collected around the 65 plants genotyped above (two 'population subsets') at the end of the dry season (between April and May of 2018). For each individual plant, we collected a soil sample proximal to the root (at a lineal distance between 0 and 7 cm) for chemical analysis. The four soil samples from each plant were combined in a composite sample in the field. Composite soil samples were air dried (between 25 and 30 ºC) for 48 hours, to prevent certain components from being lost at higher temperatures, such as those from ovens (Ren et al., 2016; Zhou et al., 2016). Soils were then sieved through a 2-mm mesh to eliminate large rocks and roots, and analyzed in the laboratory. This helped to visually determine physical differences associated with the intervened soil and the nursery substrate for reforested plants and the soil around naturally regenerated individuals.

We determined soil pH of the composite aliquot after suspending samples in deionized water (1:2.5 w/v) and shaking them for 10 min (Ren et al., 2016; Chen et al., 2019). We determined the organic carbon (C, mg C/g), total nitrogen (TN, mg N/g) and total phosphorus (TP, mg P/g)

concentrations by automatized methods (Anderson and Ingram, 1993; Roa-Fuentes et al., 2015). We additionally assessed the concentration of mineral N in form of ammonium (NH_4^+, µg/g) and nitrate (NO_3^-, µg/g) after extraction with KCl (Robertson et al., 1999). We used 0.1 g of dried green needles from a composite sample to quantify N (N_{leaf}) and P (P_{leaf}) concentrations, using Kjeldahl's method (Anderson and Ingram, 1993; Sáez-Plaza et al., 2013). Finally, we calculated soil C:TN, TN:TP, and NO_3^-:NH_4^+ ratios, and N_{leaf}: P_{leaf} ratio.

2.2.4 Statistical analysis

2.2.5.1 Phenotypic traits

Differences between groups of individuals (i.e., NR and RF) were determined for each phenotypic and soil trait through *t*-tests at 5 % probability levels. Correlations between pairs of phenotypic or soil traits were then performed with the *corrplot* package (Wickham, 2019) in R v.3.4.4 (R Core Team, 2017) to exclude highly correlated variables (Pearson's $|r| \geq 0.75$; Fig. S1). According to Gienapp et al. (2019) and Mahony et al. (2019), we fitted the following model for each *i* individual, by taking into account its age, and location within the stand (spatial effect) to eliminate experimental effects, and 'precorrect' phenotypes (6):

$$y_{ij} = \mu + age_{i,j} + loc_i + \varepsilon_i \quad (6)$$

where y_{ij} is the phenotype of individual *i* from in origin *j* (NR or RF), μ is the experimental mean, age_i the fixed age effect (as a factor), loc_i is the random and non-genetic location effect ("environment effect") and ε_i is the residual error. We report phenotypes as *z*-standardized residuals of a linear mixed effects model, as implemented in *ASReml* package (Butler et al. 2007).

We performed a PCA on 'precorrected' phenotypes using uncorrelated traits with the *factoextra* package (Kassambara and Mundt, 2020) to identify differences between groups of individuals and illustrate patterns of phenotypic differentiation. Then, a permutational multivariate analysis of variance (PERMANOVA) was used to test the null hypotheses of no phenotypic differences between groups (Anderson, 2017). The PERMANOVA was based on Bray-Curtis distances (Anderson and Santana-Garcon, 2015) and performed in *vegan* version 2.3-5 (Oksanen et al., 2019).

2.2.5.2 Population genetic structure and individual relationship

The retained panel of 1,749 SNPs that was successfully genotyped for all 65 samples was used to estimate observed heterozygosity and genomic inbreeding coefficients (*fi* = *F*) in Plink v1.04 (Purcell, 2010). Population structure was inferred with a principal components analysis (PCA) with the AGHmatrix package (Amadeu et al., 2016), and the Bayesian method available in Admixture (Alexander and Lange, 2015). First, all SNPs were used to calculate a genomic relationship matrix (GRM) between individuals, which is equivalent to the formula described by VanRaden (2008). To test for population structure, the GRM matrix was spectrally decomposed with the *eigen()* function in R. Spectral decomposition of this matrix made it possible to estimate the variation captured by genomic relationship eigenvectors (GREs), which helped determining the extent of population structure between groups (Stanton-geddes et al., 2013). The first three eigenvalues, which together explained 10.33% of the genetic variance among individuals and groups of trees, were conserved and used to control for population structure in further partial canonical correspondence analysis analyses (pCCA, see below). F_{ST} (Weir and Cockerham, 1984) between NR and RF were calculated using the "*genet.dist()*" function in *hierfstat* (Goudet and Jombart, 2015). Admixture was used to probabilistically assign individuals to a pre-defined number of clusters (*K*-value) (Skotte et al., 2013). Ten independent runs were performed for *K*-values ranging from one to five. Cross-validation was used to identify the most likely *K*-value (Alexander et al., 2009; Fatokun et al., 2018). Individual PC-loads (GRE_1) were then correlated and compared with *Q-values* and phenotypic PC-loads (PC_1).

2.2.5.3 Quantitative genomic analyses

We determined the relative contribution of genetic variables (i.e., GREs) on three phenotypic datasets (i.e., growth ($Model_{G1}$), physiological ($Model_{G2}$), and all traits ($Model_{G3}$), each one composed by multiple variables (Talbot et al., 2017), with three independent Redundancy analyses (RDA) performed with the *rda* function in *vegan*, v. 2.3-5 (Oksanen et al., 2019). RDA accounts for multiple response variables allowing to determine the effect of the total genotypic variation on each individual phenotypic trait, and in all traits as a whole. We performed RDAs in two-steps, so that genomic and phenotypic databases were analyzed using multivariate linear regressions to produce multiple matrices of fitted values (Nadeau et al., 2016; Talbot et al., 2017). Then, we used the PCA of fitted values to produce canonical/constrained axes, which are linear combinations of predictors (Legendre and Legendre, 2012; Forester et al., 2018). We retained

individual loads for significant eigenvalues (*P*-value< 0.05) and used them in subsequent analyzes. Significance for all RDA's models was evaluated using 1000 runs of Monte Carlo permutations. The best model was identified through forward selection of explanatory variables, with 999 permutations and a *P*-values of 0.05, on an adjusted coefficient of determination (R^2; Peres-Neto et al., 2006).

2.2.5.4 Genotypic - environment interaction (G × E)

Using the three models above, we conducted a pCCA (Ter Braak and Verdonschot, 1995) to estimate the relative contribution of genetic structure (i.e. individual loads of genomic relationship eigenvalues; GRE_1, GRE_2 and GRE_3) and soil variables (pH, TN, TP, NO_3^-, NH_4^+, NO_3^-:NH_4^+, C, C:TN and TN:TP) to phenotypic differences between groups of individuals. We used a forward selection of variables (genomic and edaphic) with 999 Monte-Carlo permutations and *P*-values of 0.05 to identify the best model (Ter Braak and Verdonschot, 1995); we tested for statistical significance of individual loads using an ANOVA on pCCAs eigevalue (Delgado de la Flor et al., 2017).

It must be noted that planting may have substantially modified the soil around reforested plants, and indirectly bias the performance and fit of the model above. To evaluate this, we built a new series of models for predicting the first three vectors of the RDA above (which summarize the relationship between the genomic and phenotypic variances) with the top soil variables retained in the pCCA using three strategies. First, we trained the model with the NR plants and used it to predict RDA vectors for RF individuals ('across-groups'). A good model fit would indicate that soil is equivalent in both groups and that planting may have no effect on the results above. Second, we randomly selected 50% and 75% of individuals (regardless of their group, NR or RF) to train the model and used the rest of the sample to perform predictions ('Resampling'). A better fit of this model would indicate that soil is not equivalent in both groups and that planting might be affecting results, therefore influencing the variation of the phenotype. If that were the case, we built a third model ("Compound"), which was trained with all NR plants and 50% of the RF individuals, and used to predict RDA vectors for the remaining RF plants. If planting was affecting our results, this model should also perform better than the first one. All models were validated with10-fold cross-validation (CV), after splitting data using function *createFolds* in R package *caret*. Sizes of individuals in each training and validation dataset are presented in Table

S2. Following to Ma et al. (2019) the model performance was evaluated by comparing the coefficient of determination of the calibration dataset (R^2_{Cal}) with the mean R^2 of cross-validations (R^2_{CV}), and by estimating normalized mean square errors ($nRMSE_{CV}$; RMSE divided by the sample mean).

2.2.6. Genome scans for detecting candidate genes

Genotype-environment associations were tested using the latent factor mixed model available in *lfmm* package, version 1.2 (Frichot et al., 2013) for R v.3.4.4 (R Core Team, 2017) to predict adaptive genotypes (Carvalho et al., 2021). Such model allows testing for linear relationships between individual genotypes and environmental (i.e. edaphic) traits, while accounting for neutral population structure as latent factors (Capblancq et al., 2018; Rellstab et al., 2017). Following previous clustering analyses (and preliminary *lfmm* runs), we set the number of factors (K) to 2, and performed 10 runs of the algorithm. We only considered the previously selected soil characteristics (from the pCCA models) as those that should be contributing the most to local adaptation. The z-values obtained across runs were combined using the Stouffer's method (Stouffer et al., 1949). Then, a genomic inflation factor (GIF, a scale factor used to correct for deviations of P-values that do not follow a uniform distribution, Storey and Tibshirani, 2003) was computed and used to correct P-values. The distribution of P-values was further transformed into q-values using Benjamini and Hochberg approach' (1995) for False Discovery Rate (FDR; François et al., 2016). Significant associations were inferred above thresholds of 0.05 and 0.01; q-values were graphed in Manhattan plots using R's package *qqman* to ease visualization (Turner, 2018).

Finally, the flanking ≈80 bp on each side of the identified candidates were blasted for nucleotide similarity against the *TodoFirGene* database (http://plantomics.mind.meiji.ac.jp/todomatsu/), which contains an annotated gene catalog for *Abies sachalinensis* (Ueno et al., 2018). We considered that homology was significant when the maximum 'bit score' was above 100, the percentage of identity higher than 90% and the E-value below $1E^{-10}$. The most significant hits were further scored for biochemical function, based on putative orthologous similarity against the *InterproScan* database (https://www.ebi.ac.uk/interpro/search/sequence-search) (Urrestarazu et al., 2017).

2.3 RESULTS

2.3.1. Phenotypic and edaphic differences between groups of individuals

Naturally regenerated (NR) and introduced (RF) individuals showed significant differences at ten out of twelve growth traits (TH, FH, SD, BD, AGB, CR, G_{16}, WD, AG, and AGV; *P*-value < 0.05), and at three out of seven physiological traits (Ψ, RWC and P_{Leaf} ; Table S2; *P*-value < 0.05). For growth traits, RF plants had lower means than NR saplings. For physiological traits, RF plants appeared more drought-stressed at the time of the study (lower means for Ψ and RWC; -0.47 ± 0.02 MPa and 65.7 ± 1.4 %, respectively) and had higher leaf-P concentration (P_{leaf} = 1.81 ± 0.09 mg P/ g leaf) than NR individuals (Ψ = -0.32 ± 0.015 MPa; RWC = 74.6 ± 1.9 %; P_{leaf} = 1.28 ± 0.1 mg P/ g leaf). There were, however, significant correlations between various growth and physiological traits (Fig. S1), and only ten variables were retained for subsequent analyses (six growth and four physiological traits), namely: TH, SD, CR, G_{16}, L_{leaf}, WD, RWC SLA, Ψ and P_{leaf}, (Fig S1, *P*-value < 0.05).

A joint PCA for the traits above (phenotypic variation *per se*), through a PCA, showed that both groups of individuals could be visually separated with relative ease (Fig. S2; Table S3). A PERMANOVA based on Bray-Curtis distances showed that these groups are statistically different (*P*-value < 0.05, Table S4). The first two axes of the PCA explained 54.36% of the total variation and were respectively loaded by positive correlations with TH, SD, CR, Ψ and G_{16} (PC_1) and SLA (PC_2). PC_2 was also loaded by negative correlations with RWC and WD (Table S3). NR individuals had a higher phenotypic breadth than RF plants, which tended to be more phenotypically homogeneous (Fig. S2) and to have a similar age.

Soil around NR and RF plants showed significant differences at four out of nine edaphic characteristics (pH, C, TN, and NH_4^+; Table 1, *P*-value < 0.05). Soils under RF had lower organic C and TN concentrations, but greater pH and NH_4^+ concentration than those under NR. Significant co-variation was observed between some of these soil characteristics (Fig. S3, *P*-value < 0.05), which resulted in the elimination of organic C, and NO_3^- metrics from further analyses.

2.3.2 Correlation between Genetic Structure and Phenotypic Variation

Sequencing provided an average of 2,803,267 raw reads per individual. After filtering, 31,462 consensus reads were retained, which were *de novo*-assembled for identifying 373,267

SNPs with an average depth of 8X. After the final quality-filtering, the final data set included 1,749 SNPs genotyped for all 65 individuals (34 NR and 31 RF).

Five-fold cross-validation of non-partitioned genomic data in Admixture revealed an optimal number of two genetic groups (*K*-value = 2; Fig. 2A; Table S6). However, a substantial degree of admixture was observed for NR individuals and, to a lower extent, for RF plants; indicating only partial genetic differentiation. An unsupervised demographic clustering (PCA) showed similar results, with spectral decomposition of the first two genomic eigenvalues ('GRE axes' from now on) explaining less than 8% of the genomic variance (4.21 and 3.25% for GRE_1 and GRE_2, respectively; Fig. S4A). Both groups had similar values (*P*-value > 0.05) of observed heterozygosity (0.235 ± 0.007 and 0.219 ± 0.005 for NR and RF, respectively) (Fig. S4B), a similar spectrum of genetic relationships (Fig. S5), and of genomic inbreeding coefficients (0.021 ± 0.027 and 0.026 ± 0.025 for NR and RF, respectively). F_{ST} between origins was low (0.012; *P*-value = 0.62).

There was a significant correlation between individual loads for phenotypic and genetic PCs (R^2_{adj} = 0.22; *P*-value < 0.001; Fig. 2B). However, redundancy analysis (RDA), performed on different partitions of phenotypic traits (models), showed that only modest parts of the phenotypic (constrained) variance (TVE) between groups of individuals could be accounted for by genomic data (Table 2). The best model ($Model_{G3}$) included all phenotypic traits and explained ~17% of the phenotypic variance. For the two other models (growth traits only, and physiological traits only), genomic data explained 10.91 and 6.86% of the phenotypic variance, respectively (Table 2). For the best model, the first three eigenvalues contributed approximately 40% of the TVE (RDA_1 = 16.88%, RDA_2 = 12.07%, RDA_3 = 11.35%; Fig. S6). They were mostly loaded by G_{16}, Ψ, TH and CR (RDA_1), SD, WD and L_{leaf} for RDA_2, variables and RWC, SLA and P_{leaf}, for (RDA_3), respectively. These three axes were retained for analyses below.

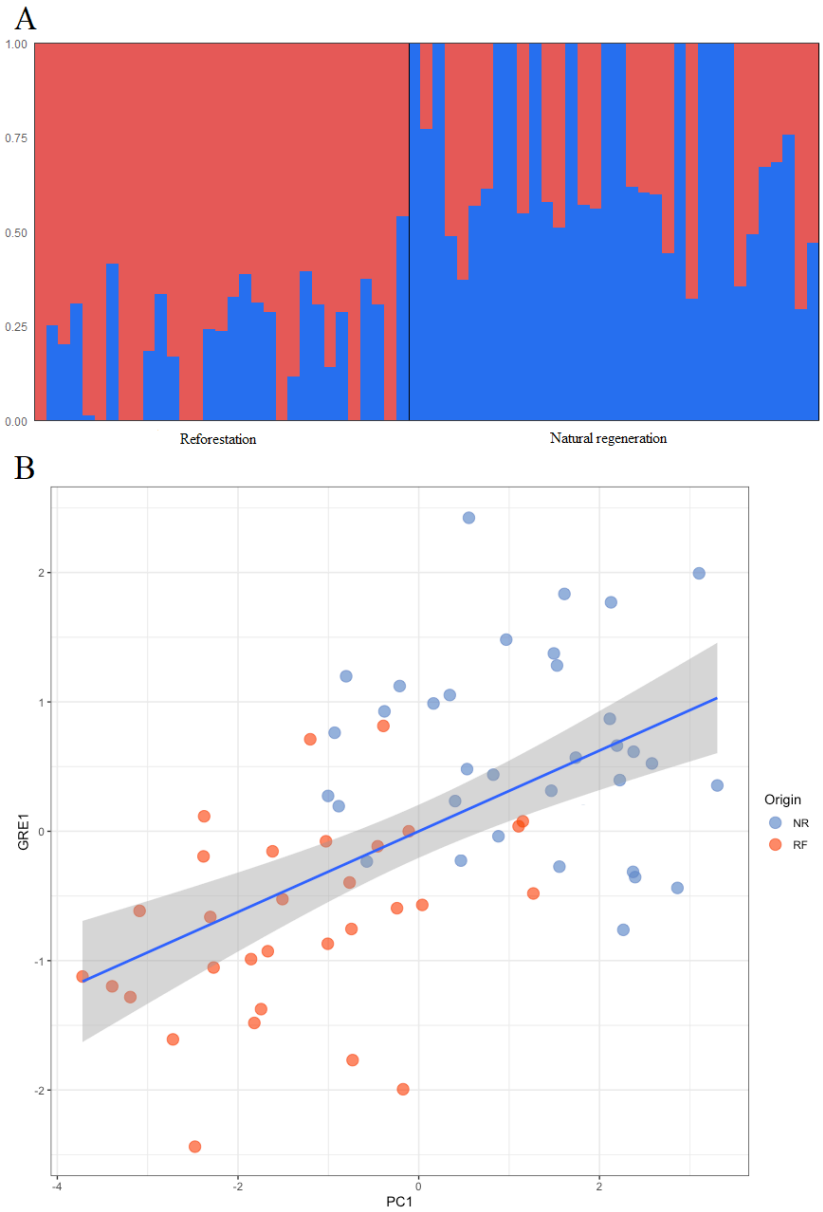

Fig 2. A Genomic differentiation (barplot) among Sacred fir individuals sampled in a trial site in Central Mexico. An Admixture clustering was performed with 1,749 SNPs and assuming *K-value* = 2 (lowest cross-validation error). **B** Correlation between the phenotypic (PC1, x axis) and genetic (GRE1, y axis) principal components grouped by individual class (management). The diagonal line represents the best fit regression line (R^2_{adj} = 0.22; *P*-value = 4.2 E^{-5}). Blue and red colors represent naturally regenerated (NR) and introduced reforested (RF) individuals, respectively.

3.3.3 Including soil characteristics for partitioning phenotypic variation with genomic data

To specifically display the effect of within-population genotype × environment (G × E) on phenotypic variation, we performed pCCA ordinations with an interactive forward selection of explanatory variables. In all cases, the best models comprised four soil biogeochemical metrics

(TP, TN and NH_4^+ concentrations, and NO_3^-:NH_4^+ ratio) and the two genomic eigenvalues obtained above (GRE_1 and GRE_2). Integrating genome-soil interactions had a significant impact on TVE for all three groups of phenotypic traits (5.63, 3.79 and 6.88%, for $Model_{G\times E1}$, $Model_{G\times E2}$ and $Model_{G\times E3}$, respectively), when compared to models exclusively using genomic or soil data alone (Table 3; Table S7). These models were also better (lower *P*-value) and had higher TVE values than the previous RDAs. As for previous analyses, the model including all phenotypic traits ($Model_{G\times E3}$) had the highest TVE (21.58%). The first two eigenvalues for this model explained up to 94% of the TVE (Fig. 3; Table S8), and were mostly loaded by soil TP, TN and NH_4^+, and GRE_1 (pCCA1), and by soil NO_3^-:NH_4^+ ratio and GRE_2 (pCCA2; Table S9 and S10), respectively. The relationship between RDA_1, and soil TN and NH_4^+ concentrations (*P*-value < 0.01) shown in Fig. 4 illustrates how integrating soil variation can account for a higher proportion of the phenotypic differentiation between plant groups than using genomic data alone.

As a proxy for determining whether soil differences between groups were affecting results, we built a series of models for predicting RDA vectors with the top soil traits loading the pCCA axes above. The model built 'across-groups' (NR → RF) only explained a small percentage of the variance of the calibration dataset for all RDA vectors (Table 4). Model performance improved when including RF plants into the training set ('Compound' in Table 4) and when using a random set of individuals from both groups ('Resampling') to calibrate the model. The best fit was achieved when the 'Resampling' was randomly sampled 75% as training set (Table 4). Thus indicating that soil differences between groups have a measurable effect on phenotype and that planting might be affecting individual performance.

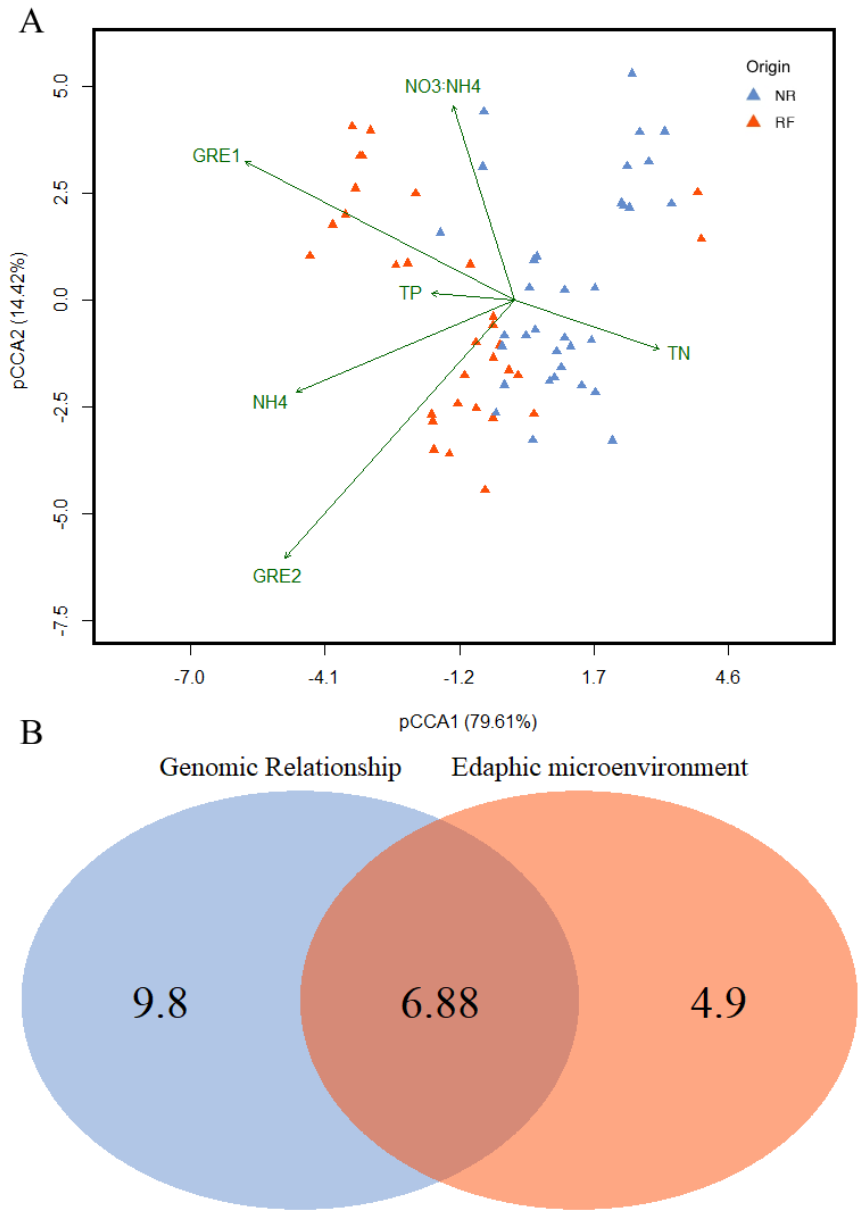

Fig. 3. A. Biplot derived from the partial canonical correspondence analysis for $Model_{GxE3}$ (all phenotypic traits included) according to edaphic characteristics in a Sacred fir trial site in Central Mexico. TN, total soil nitrogen; TP, total soil phosphorous; NH_4^+, mineral soil nitrogen in form of ammonium; NO_3^-:NH_4^+ ratio of mineral soil nitrogen in form of nitrate and ammonium; GRE_1 and GRE_2, genomic relatedness eigenvalues among individuals. **B.** Percentage of the total phenotypic variance between groups of individuals (NR vs. RF) explained by each database in the partial canonical correspondence analysis (pCCA; see **Table 3**).

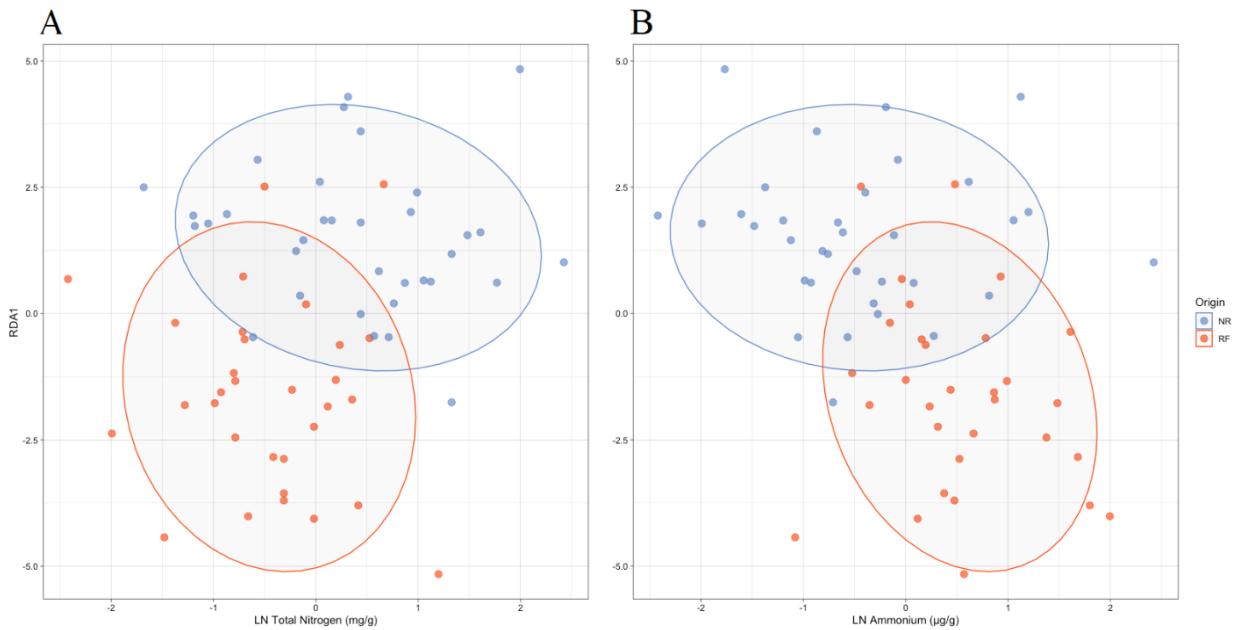

Fig. 4. Biplot of the most significant component of a redundancy analysis for $Model_{G\times E3}$ between phenotypic and genotypic data for two groups of individuals (NR and RF) of Sacred fir sampled in a trial site in Central Mexico. **A.** Total soil nitrogen (TN, mg N/g DW) and **B.** mineral soil nitrogen in form of ammonium (NH_4^+, µg/g). NR, naturally regenerated individuals; RF, introduced reforested plants. There were significant differences between the 'populations' centroids for both edaphic characteristics according to a PERMANOVA (*P*-value < 0.01).

2.3.4. Genotype-environment associations (GEA)

Testing for associations between SNPs and soil characteristics (by including population structure as a latent factor, $K = 2$, and using strict GIF and FDR corrections; Fig. 5) revealed fourteen candidate loci (0.48% of the total; Table S11**).** Half of them (7) were associated with the concentration of TN in the soil, five were correlated to the ratio of both inorganic soil N forms (i.e., NO_3^-:NH_4^+ratio), and the remaining two were associated with the concentration of NH_4^+. Blasting contig sequences that contained these outliers against the transcriptome of *A. sachalinensis* allowed annotating two regions, both with SNPs associated to soil TN (Table S11). The first one was similar to a member of the *pentatricopeptide repeat protein* (*PPR*) superfamily (and thus related to protein amino acid binding; Barkan and Small, 2014), and the second one to a *glycosyl hydrolase 9B13* gene (*GH9B13*; related to catalytic activities in carbohydrate metabolism; Opassiri et al., 2006). Genotype-environment associations suggested a dominance model for these two loci (Fig. S7).

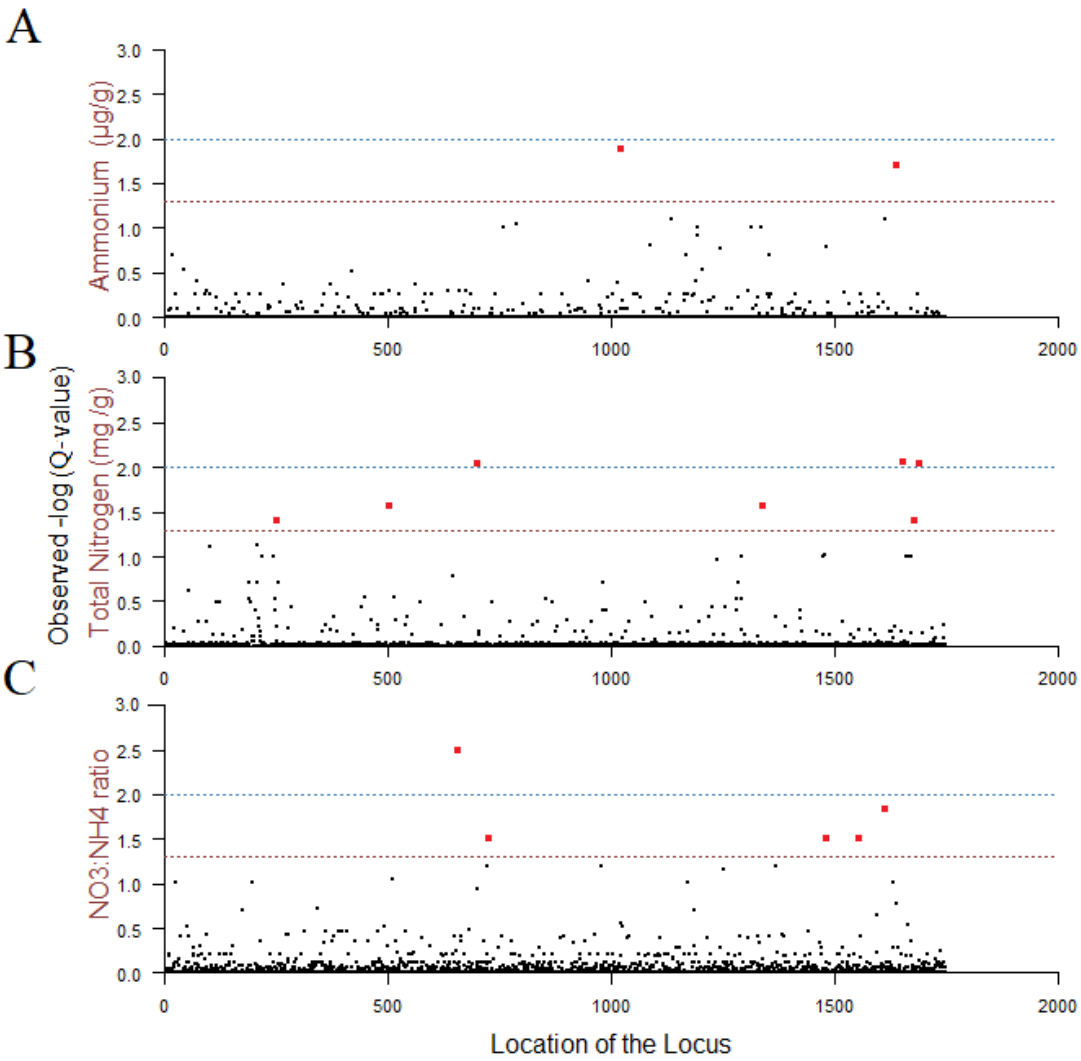

Fig 5 Manhattan plots derived from latent factors mixed models testing for genotype-edaphic associations in a Sacred fir trial site in Central Mexico. Plots represent the −log10 significance values (*Q*-value) obtained for 1,749 SNPs. **A**, mineral soil nitrogen in form of ammonium (NH_4^+): 2 SNPs; **B**, total soil nitrogen (TN): 7 SNPs; **C**, ratio of mineral soil nitrogen in form of nitrate and ammonium (NO_3^-:NH_4^+): 5 SNPs. Red and blue lines indicate FDR values of 0.05 and 0.01, respectively. The largest red dots show the most significant associations between genetic and microenvironmental variables.

2.4. DISCUSSION

The present study provides a quantitative overview of the interaction between genotype and soil heterogeneity and its effects on plant phenotype at the local scale. It also shows that taking soil-plant interactions, and planting conditions (soil disturbance), into account can enhance phenotypic predictability in reforestation and restoration plans. It further reports two novel candidate genes associated with soil N availability, which deserve additional investigation under controlled conditions and in an evolutionary context.

2.4.1. Correlation between phenotypic and genetic variation in natural test-trial

Understanding species' phenotypic variability requires the identification of evolutionary, demographic and ecological processes operating at the population level (Mitchell-Olds et al., 2007; Villellas et al., 2014). In small and secluded natural populations or plantations, phenotypic performance and sapling survival can be compromised because of reduced genetic diversity, such as suspected herein from the differences between groups of individuals (lower growth and more stressed plants for the RF than for the NR group; Table S3). However, given that edaphic, phenotypic and genomic variability allow to subtly differentiate both types of trees, it seems necessary to invoke other forces affecting phenotypic performance.

Genetic differentiation was weak between NR and RF individuals (F_{ST} =0.012 and *P*-value =0.62), especially when compared to the range-wide population levels (G_{ST} =0.20-0.51; Cruz-Nicolás et al., 2020), which, together with Admixture and PCA results, suggests a regional origin for the reforested individuals (see also Arenas et al., 2021). This indicates that even when privileging local seed-sources for reforestation, to avoid introducing maladapted individuals, phenotypic differences can still occur between reforested and regenerated plants, even ~9 years after plantation. A possible explanation for this could be translocation-related effects, like those derived from soil layer mixing and compaction during plantation (Lawson and Michler, 2014; Sebastian-Azcona et al., 2020). However, for adequately evaluating planting effects in forest restoration, requires a different experimental approach, as for example, including several repetitions at various locations with contrasting soil and topographic conditions.

Other than translocation effects, our results also hint that phenotypic variation might be accounted for, at least in part, by genetic factors. An initial indication is the correlation between phenotypic traits (i.e. those associated to PC1: TH, SD, CR, Ψ, and G_{16}) and the first axis of genetic differentiation (GRE_1; R^2_{adj} = 0.22). However, such a low coefficient is likely influenced by environmental variation (Forester et al., 2018) and probably includes plastic and epigenetic effects (Wang et al., 2020). While some studies have previously observed a similar correspondence between genetic and phenotypic differentiation in forest plants (e.g., Eckert et al., 2013; Gömöry et al., 2013; Harter et al., 2015), molecular markers are generally considered poor predictors of adaptive genetic variability (Kozak et al., 2011), unless they are massively surveyed and integrated into sophisticated genome-prediction models (Chen et al., 2018; Lenz et al., 2020). Herein, simple RDA models accounted for a modest ~17% of the phenotypic difference between

groups when including all retained traits (Table 2); thus suggesting that models with both more markers (especially from coding genes; Lebedev et al., 2020) and individuals could explain larger parts of the phenotypic variance, especially if they incorporate factors like G × E or soil variability.

2.4.2 Significant soil edaphic variation at small spatial scales

Before being integrated into predictive models, environmental variability had to be described. Significant soil variation was detected along the small spatial gradient surveyed in the study site, with differences observed for soil nutrient concentrations and ratios (e.g., TN concentration and NO_3^-:NH_4^+ ratio), which were particularly important for explaining differences between plant groups. Putative explanations for such soil variability include the above-mentioned planting-effects (see also Sauer et al., 2012; Larkin et al., 2016; Chen et al., 2019), together with other plant-soil feedbacks. These may influence nutrient availability via nutrient deposition and mineralization, which modify soil nutrient content at short spatial scales (Wardle et al., 2004; Silva and Lambers, 2020). Other anthropic factors derived from forest management, like machinery use during wood extraction, could also have changed the soil physical properties and might also be at stake (Cambi et al., 2015; Elliot et al., 2019). Reforestation usually includes removing the upper organic layer, and planting seedlings along with nursery soil altering the upper soil layers, and consequently the soil biogeochemical properties and nutrient cycles at reforested points (Le et al., 2012). Altogether, this requires translocated individuals to rapidly adapt to novel soil selective pressures during the acclimation time, which may diminish growth and nutrient assimilation, particularly N (Lu et al., 2012; Liu et al., 2019). In our trial site, increased soil NH_4^+ concentrations were estimated for reforested plants, perhaps from an exogenous source (i.e. nursery soil was enriched with N fertilizers; Andrés Juárez pers. comm.), while higher soil TN and organic carbon concentrations were determined around NR plants (Table 1 and Fig. S3). Such differences suggest higher organic matter inputs for the natural seedlings (Lafleur et al., 2015), which could increase soil water retention and have measurable effects in plant phenotypic performance. A better understanding of microspatial variability of soil fertility seem thus important to predict plant functional traits (Li et al., 2018), adequate reforestation practices and improve their success (Koskela et al., 2014; Larkin et al., 2016).

2.3.3. Partitioning phenotypic variation due to genotype by microenvironmental factors and their interaction

Our results highlight the potential that models combining growth and physiological data may have for describing individuals' fitness in natural forest tree populations. It has been well established that plant growth, and other factors such as photosynthetic activity and leaf nutrients (Čepl et al., 2016), have a genetic basis that can be influenced by the environment, including soil conditions (e.g., Lojewski et al., 2009; Madritch and Lindroth, 2011). Thus, identifying the effects of these conditions and their interaction with individual's genotypes should help increasing phenotypic predictability, such as shown herein (Tables 4 and S7). Indeed, models including only genotypic data accounted for less than 17% of the phenotypic variance, while models exclusively taking soil data into account only explained ~12%. In contrast, models including both factors and their interaction accounted for up to about 22% in TVE,suggesting a local phenotypic response that depends on soil nutrient content and within-population G × E (which includes planting conditions). Our results are consistent with those from common gardens in other conifers, which report a significant positive correlation between plant growth and soil N concentrations (Deng et al., 2019; Ren et al., 2016), and support the view that terrestrial vegetation growth is generally N-limited (Augusto et al. 2017; Du et al., 2020). This indicates that local polygenic adaptation, quick response to a new environment, and phenotypic plasticity could be a driving force for seedling establishment in managed forest populations.

Testing for local adaptation requires, however, estimating fitness in common garden or reciprocal transplant experiments (Blanquart et al., 2013; Savolainen et al., 2013; Lascoux et al., 2016), including the comparison with traits under strong genetic control, such as bud set timing (Mckown et al., 2014; Abbot et al., 2015), and evaluating survival over time (Cruzado-Vargas et al., 2019). Our use of individuals from two origins exposed to similar environmental pressures for over nine years is a first step for addressing microgeographic adaptation. Heritability estimates for some of the studied traits (h^2 between 0.12 - 0.29; Arenas et al., 2021) indicate that there is a considerable amount of additive variation in these traits upon which natural selection could act (Ramírez-Valiente et al., 2014; Tiffin and Ross-Ibarra, 2014), and our predictive models suggested differential phenotypic response between origins (Tables 3 and 4). Altogether, these results indicate an interaction between plant genotype and microenvironment that can be measured even after nine years of planting.

Spatially variable selection has been reported in other tropical trees (Brousseau et al., 2020; Carvalho et al., 2021). Still, the small number of trees tested, the limited microenvironmental information (soil studied in a single "snap-shot" in time and the edaphic range was rather limited), and the lack of evidence for phenotypic plasticity hamper us, from extrapolating our results to larger spatial or temporal scales. To detect local adaptation, such scales have to be taken into account, as the intensity and direction of selection is likely to vary over space and time (Brachi et al., 2013; de Villemereuil et al., 2018). In forest trees, different life stages may have different fitness optima, given than environmental conditions change within and among years (Scotti et al., 2016).

As with previous studies in controlled and natural conditions (Talbot et al., 2017; Budde et al., 2013; Riordan et al., 2016), a large portion of the phenotypic variance remained unaccounted for in our study. Usual explanations for this include polygenic adaptation, epigenetic factors, planting effects, and unmeasured (or inaccuracy in measured) environmental traits that could be playing important roles in phenotypic performance (Arnold et al., 2019; Des Marais et al., 2013; Talbot et al., 2017). Biotic interactions such as those with mycorrhiza, soil bacterial communities, herbivores or nurse plants have all been shown to be important for plant development and survival in shade-tolerant plants like sacred fir (Bennett and Klironomos, 2019; Carbajal-Navarro et al., 2019; Liu et al., 2019); taking these factors into account might help increasing TVE. Here, we made sure of including only individuals without visible herbivore damage and all growing under similar light conditions, leaving the effect of planting, including modified soil microbiome composition (i.e., bacteria and mycorrhizae), as the most likely unaccounted factor. For instance, the apparently higher P uptake that had the RF plants, evidenced by their increased P_{leaf} when compared to NR plants, could be a consequence of either the soil or its microbiome composition; more specifically of the mycorrhizae inoculated during the nursery phase (Andrés Juárez pers. comm.). While this is still a hypothesis to test, multiple studies suggest that the presence of bacterial and mycorrhizal communities can alter the phenotypic performance in a wide range of plant traits (Partida-Martínez and Heil, 2011), even at the biochemical and anatomical levels, by favoring root N and P absorption. Therefore, the correlation between the measured phenotypic traits and soil NH_4^+ and TN concentrations could be a consequence of reforestation, even more than nine years after planting. This is certainly a worth exploring avenue that may require testing in multiple sites with controlled soil NH_4^+ and TN concentrations, and using incubation

techniques.

2.4.4. Soil microenvironment: a potential selective factor for Sacred fir

Our results also highlight the importance of studying soil fine-scale variation and its potential effects on phenotype in more detail. Indeed, in spite of using a (very) reduced representation of an extremely large genome (~18 Gb; Mosca et al., 2019), we detected 14 candidate loci potentially related to soil variation, and were able to annotate two of them (Table S11). One of these genes is a member of the *PPR* gene superfamily, which have been related to drought stress responses, and molecular mechanisms like protein binding (Barkan and Small, 2014). A member of this superfamily was previously related to adaptation to soil conditions and drought in *Medicago truncatula* (Guerrero et al., 2018). The other annotated gene, *GH9B13*, has been related to cellulose biosynthesis (Opassiri et al., 2006), and has been shown to recurrently over-express in trees under severe drought (Spokevicius et al 2017). Interestingly, this gene is also induced upon endophytic bacterial colonization in *Arabidopsis* (Sheoran et al. 2016); thus, endorsing its potentially adaptive role to soil composition. Future works under controlled conditions (e.g. incubation experiments) that include a better representation of the annual microenvironmental variation, and using a larger sample size are thus needed to confirm the putatively adaptive roles of these genes.

Finding potential candidate genes to soil adaptation is not new in tree species, including Sacred fir. For instance, Méndez-González et al. (2017) found several genes related to soil fertility, which accounted for a significant portion of the within-population structure in a Sacred fir stand located some kilometers south from our trial site. Our results further reinforce previous findings (but a smaller scale) that suggested soil adaptation in *Pinus contorta* (Eckert et al., 2012), where a series of candidate genes (e.g., ABC-Transporters) associated with gradients of soil aluminum and phosphate concentrations, (i.e., soil toxicity and nutrient supply) were detected. Such predictive models could be helpful for supporting future forest management and assisted migration plans (Sáenz-Romero et al., 2016; Aitken and Bemmels, 2016; George et al., 2017).

2.4.5. Tree genotypes and interactions with soil properties and management implications

Seedlings are exposed to a range of biotic and abiotic pressures, for instance competition, herbivory, drought and disturbance. Theoretical models suggest that polygenic adaptation may account for a large portion of phenotypic diversity between populations, which should be maintained by a combination of natural selection and gene flow (Le Corre and Kremer, 2012), and perhaps reinforced by demographic or neutral processes. Our pilot study points that genotype, planting conditions and soil variation may interact and impact tree phenotypic performance at small geographic scales (i.e. tens of meters). Bonifying these models with novel phenotyping capabilities and genomic resources will allow for a finer selection of genomic variants, and help identifying additional links between genetic divergence and environment. Expanding these studies to other populations along the species' natural range should be of particular interest, since selection is likely to target multilocus combinations, rather than particular polymorphisms (Scotti et al., 2016), which may change among populations and environmental conditions.

The link between plant genotype and soil conditions are poorly understood, both in controlled and natural conditions (Beals et al., 2020). Individual tree genotypes and genotype diversity play crucial roles in shaping community structure, and their interaction in a changing environment may modify ecosystem function (Arnold et al., 2019), and composition (Purahong et al., 2016). The interaction between particular genotypes and mutualistic communities may improve tree survival, growth and physiological performance, which is crucial to predict, given the forecasted threats associated to future climate change (Gehring et al., 2017). Although potentially difficult to address, including within-population G × E and polygenic adaptation should thus be critical for guiding reforestation strategies (Aitken and Whitlock, 2013). For instance, for optimizing planting conditions and avoiding maladapted individuals to particular conditions (Isabel et al., 2020).

2.5. REFERENCES

Abdi, E., Moghadamirad, M., Hayati, E., Jaeger, D., 2017. Soil hydrophysical degradation associated with forest operations. Forest Science and Technology 13, 152–157. https://doi:10.1080/21580103.2017.1387611.

Abbott, A.G., Zhebentyayeva, T., Barakat, A., Liu, Z., 2015. The genetic control of bud-break in trees. Adv. Bot. Res. 74, 201–228. https://doi: 10.1016/bs.abr.2015.04.002

Abràmoff, M.D., Magalhães, P.J., Ram, S.J., 2005. Image processing with ImageJ Part II. Biophotonics International 11, 36–43. https://doi:10.1117/1.3589100.

Aitken, S.N., Bemmels, J.B., 2016. Time to get moving: assisted gene flow of forest trees. Evolutionary Applications 9, 271– 290. https://doi.org/10.1111/eva.12293.

Aitken, S. N., Whitlock, M. C., 2013. Assisted gene flow to facilitate local adaptation to climate change. Annual Review of Ecology, Evolution, and Systematics 44, 367–388. https://doi:10.1146/annurev-ecolsys-110512-135747.

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

Alexander, D.H., Lange, K., 2015. Admixture 1.3 Software Manual. 3–4.

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Research 19, 1655–1664. https://doi:10.1101/gr.094052.109.

Amadeu, R., Cellon, C., Lara, L., Resende, M., Oliveira, I., Ferrao, L., Munoz, P., Garcia, A., 2019. Package 'AGHmatrix' 1.0.2. Relationship matrices for diploid and autopolyploid species.

Anderson, J.M., Ingram, J.S.I., 1993. Tropical soil biology and fertility: A Handbook of Methods. The Journal of Ecology 78, 547. https://doi:10.2307/2261129.

Anderson, M.J., 2017. Permutational multivariate analysis of variance (PERMANOVA). Wiley StatsRef: Statistics Reference Online 1–15. https://doi:10.1002/9781118445112.stat07841.

Anderson, M.J., Santana-Garcon, J., 2015. Measures of precision for dissimilarity-based multivariate analysis of ecological communities. Ecology Letters 18, 66–73. https://doi:10.1111/ele.12385.

Arnold, P.A., Kruuk, L.E.B., Nicotra, A.B., 2019. How to analyse plant phenotypic plasticity in response to a changing climate. New Phytologist 222, 1235–1241. https://doi:10.1111/nph.15656.

Arenas, S., Cortés, A.J., Mastretta-Yanes, A. Jaramillo-Correa J.P. 2021. Evaluating the accuracy of genomic prediction for the management and conservation of relictual natural tree populations. Tree Genetics & Genomes 17, 12. https://doi.org/10.1007/s11295-020-01489-1.

Augusto L., Achat D.L., Jonard M., Vidal D., Ringeval B., 2017. Soil parent material-A major driver of plant nutrient limitations in terrestrial ecosystems. Glob Chang Biol. 23 (9), 3808-3824. https://doi:10.1111/gcb.13691.

Baer, S. G. 2016. Nutrients as determinants and endpoints in ecological restoration. Pages 333-364. In M. A. Palmer, J. Zedler, and D. Falk (editors). Foundations of Restoration Ecology, 2nd edition. Island Press: Washington D.C.

Bailey, J.K., Schweitzer, J.A., Úbeda, F., Koricheva, J., LeRoy, C.J., Madritch, M.D., et al., 2009. From genes to ecosystems: A synthesis of the effects of plant genetic factors across levels of organization. Philosophical Transactions of the Royal Society B: Biological Sciences 364, 1607–1616. https://doi:10.1098/rstb.2008.0336.

Barkan, A., Small, I., 2014. Pentatricopeptide Repeat Proteins in Plants. Annual Review of Plant Biology 65, 415–442. https://doi:10.1146/annurev-arplant-050213-040159.

Beals, K.K., Moore, J.A.M., Kivlin, S.N., Bayliss, S.L.J., Lumibao, C.Y., Moorhead, L.C., Patel, M., Summer, J.L., Ware, I.M., Bailey, J.K., Schweitzer, J.A., 2020. Predicting Plant-soil feedback in the field : meta-analysis reveals that competition and environmental stress differentially influence PSF, Front. Ecol. Evol., 8 (121) https://doi:10.3389/fevo.2020.00191.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society B 57 (1), 289–300. https://doi:10.2307/2346101.

Bennett, J. A., Klironomos, J., 2019. Mechanisms of plant–soil feedback: interactions among biotic and abiotic drivers. New Phytologist 222, 91–96. https://doi:10.1111/nph.15603.

Blanquart, F., Kaltz, O., Nuismer, S.L., Gandon, S., 2013. A practical guide to measuring local adaptation. Ecology Letters 16, 1195–1205. https://doi:10.1111/ele.12150.

Boyer, J.S., 1967. Leaf water potentials measured with a pressure chamber. Plant physiology, 42, 133–137.

Brachi, B., Villoutreix, R., Faure, N., Hautekèete, N., Piquot, Y., Pauwels, M., Roby, D., Cuguen, J., Bergelson, J., Roux F., 2013. Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in Arabidopsis thaliana. Molecular Ecology 22 (16), 4222–4240. https://doi:10.1111/mec.12396.

Brousseau L., Fine P.V.A. Dreyer E., Vendramin G.G., Scotti I., 2020. Genomic and phenotypic divergence unveil microgeographic adaptation in the Amazonian hyperdominant tree Eperua falcata Aubl. (Fabaceae). Molecular Ecology https://doi:10.1111/mec.15595.

Budde K.B., Heuertz M., Hernández-Serrano A., Pausas J.G., Vendramin G.G., Verdu 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, 230–241.

Butler D., Cullis B. R. Gilmour A. R. Gogel B. J., 2007. ASReml-R estimates variance components under a general linear mixed model by residual maximum likelihood (REML).

Cambi, M., Certini, G., Neri, F., and Marchi, E., 2015. The impact of heavy traffic on forest soils: A review. Forest Ecology and Management 338, 124–138. https://doi:10.1016/j.foreco.2014.11.022.

Cambi, M., Paffetti, D., Vettori, C., Picchio, R., Venanzi, R., Marchi, E., 2017. Assessment of the impact of forest harvesting operations on the physical parameters and microbiological components on a Mediterranean sandy soil in an Italian stone pine stand. European Journal of Forest Research 136, 205–215. https://doi:10.1007/s10342-016-1020-5.

Capblancq, T., Luu, K., Blum, M. G. B., Bazin, E., 2018. Evaluation of redundancy analysis to identify signatures of local adaptation. Molecular Ecology Resources 18, 1223–1233. https://doi:10.1111/1755-0998.12906.

Cappai C., Kemanian A.R., Lagomarsino A., Roggero P.P., Lai R., Agnelli A.E. Seddaiu G., 2017. Small-scale spatial variation of soil organic matter pools generated by cork oak trees in Mediterranean agro-silvo-pastoral systems. Geoderma 304, 59–67. https://doi:10.1016/j.geoderma.2016.07.021

Carbajal-Navarro, A., Navarro-Miranda, E., Blanco-García, A., Cruzado-Vargas, A.L., Gómez-Pineda, E., Zamora-Sánchez, C., et al., 2019. Ecological restoration of *Abies religiosa* forests using nurse plants and assisted migration in the Monarch Butterfly Biosphere Reserve, Mexico. Frontiers in Ecology and Evolution 7. https://doi:10.3389/fevo.2019.00421.

Carrasco-Carballido, V., Martínez-Garza, C., Jiménez-Hernández, H., Márquez-Torres, F., Campo, J., 2019. Effects of initial soil properties on three-year performance of six tree species in tropical dry forest restoration plantings. Forests 10, 428.

Carvalho, C.S., Forester, B.R., Mitre, S.K., Alves, R., Imperatriz-Fonseca, V.L., Ramos, S.J., Resende-Moreira, L.C., Siqueira, J.O., Trevelin, L.C., Caldeira, C.F., Gastauer, M., Jaffé, R., 2021. Combining genotype, phenotype, and environmental data to delineate site-adjusted provenance strategies for ecological restoration. Molecular Ecology and Resourses 21(1),44-58. https://doi: 10.1111/1755-0998.13191.

Čepl, J., Hola, D., Stejskal, J., Korecki, J., Kocova, M., Lhotakova, Z., et al., 2016. Genetic variability and heritability of chlorophyll a fluorescence parameters in Scots pine (Pinus sylvestris L.). Tree Physiology 36, 883–895. https://doi:10.1093/treephys/tpw028.

Chen, J., Shen, W., Xu, H., Li, Y., Luo, T., 2019. The composition of nitrogen-fixing microorganisms correlates with soil nitrogen content during reforestation: A comparison between legume and non-legume plantations. Frontiers in Microbiology 10, 1–11. https://doi:10.3389/fmicb.2019.00508.

Chen, Z.Q., Baison, J., Pan, J., Karlsson, B., Andersson, B., Westin, J., et al., 2018. Accuracy of genomic selection for growth and wood quality traits in two control-pollinated progeny trials using exome capture as the genotyping platform in Norway spruce. BMC Genomics. 19 (946). https://doi:10.1186/s12864-018-5256-y.

Chojnacky, D.C., Heath, L.S., Jenkins, J.C., 2014. Updated generalized biomass equations for North American tree species. Forestry 87, 129–151. https://doi:10.1093/forestry/cpt053.

CONAGUA, Comisión Nacional del Agua. 2019. https://smn.conagua.gob.mx/ es/informacion-climatologica-por-estado?estado=mex

Correia A.C., Mutke S., Silva J., 2017. Variability of specific needle area in Pinus pinea L. with environment resources availability: light, water and nutrients. In : Carrasquinho I. (ed.), Correia A.C. (ed.), Mutke S. (ed.). Mediterranean pine nuts from forests and plantations. Zaragoza : CIHEAM. p. 43-47 (Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 122)

Cruz-Nicolás, J., Giles-Pérez, G., Gonzáles-Linares, E., Múgica-Gallart, J., Lira-Noriega, A., Gernandt, D.S., Eguiarte L.E., Jaramillo-Correa, J.P., 2020. Contrasting evolutionary processes drive morphological and genetic differentiation in a subtropical fir (*Abies*, Pinaceae) species complex. Botanical Journal of the Linnean Society 192 (2), 401–420.

Cruz-Ruiz E., Cruz-Ruiz A., Aguilera-Gómez L.I. Norman-Mondragón H.T., Velázquez R.A., Nava-Bernal G., 2012. Efecto en las características edáficas de un bosque templado por el cambio de uso de suelo. Terra latinoamericana 1, 189–197.

Cruzado-Vargas, A.L., Zamudio-Sánchez, F.J., Rodríguez-Yam, G.A., Carbajal-Navarro, A.L., Blanco-García, J.A., Sáenz-Romero, C., 2019. Growth of naturally regenerated *Abies religiosa* (Kunth) Schltdl. and Cham. Seedlings in a nursery and genetic variation among provenances. Revista Chapingo, Serie Ciencias Forestales y del Ambiente 26, 85–96. https://doi:10.5154/r.rchscfa.2019.01.013.

de Mita, S., Thuillet, A.C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., Vigouroux Y., 2013. Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. Molecular Ecology 22 (5), 1383–1399. https://doi:10.1111/mec.12182.

de Villemereuil, P., Mouterde, M., Gaggiotti, O.E., Till-Bottraud, I., 2018. Patterns of phenotypic plasticity and local adaptation in the wide elevation range of the alpine plant. Arabis alpina. Journal of Ecology 106, 1952–1971. https://doi:10.1111/1365-2745.12955.

Delgado de la Flor, Y.A., Burkman, C.E., Eldredge, T.K., Gardiner, M.M., 2017. Patch and landscape-scale variables influence the taxonomic and functional composition of beetles in urban greenspaces. Ecosphere 8 (11) e02007. https://doi:10.1002/ecs2.2007.

Deng, S., Wipf, H.M.L., Pierroz, G., Raab, T.K., Khanna, R., Coleman-Derr, D., 2019. A plant growth-promoting microbial soil amendment dynamically alters the strawberry root bacterial microbiome. Scientific Reports 9, 1–15. https://doi:10.1038/s41598-019-53623-2.

Des Marais, D.L., Hernandez, K.M., 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, 5–29. https://doi:10.1146/annurev-ecolsys-110512-135806.

Du, E., Terrer, C., Pellegrini, A.F., Ahlström, A., van Lissa, J.C., Zhao, X., Xia, N., Wu, X., Jackson, R.B., 2020. Global patterns of terrestrial nitrogen and phosphorus limitation. Nature Geoscience 13, 221–226.

Dufour, A., Gadallah, F., Wagner, H.H., Guisan, A., Buttler, A., Dufour, A., et al., 2006. Plant species richness and environmental heterogeneity in a mountain landscape : effects of variability and spatial configuration. Ecopgraphy 29 (4), 573–584.

Eaton, D.A.R., 2014. PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics 30, 1844–1849. https://doi:10.1093/bioinformatics/btu121.

Eckert, A.J., Shahi, H., Datwyler, S.L., Neale, D.B., 2012. Spatially variable natural selection and the divergence between parapatric subspecies of lodgepole pine (Pinus contorta, Pinaceae). American Journal of Botany 99, 1323–1334. https://doi:10.3732/ajb.1200055.

Eckert, A.J., Wegrzyn, J.L., Liechty, J.D., Lee, J.M., Cumbie, W.P., Davis, J.M., Goldfarb B., Loopstra C.A., Palle S.R., Quesada T., Langley C.H., Neale D.B., 2013. The evolutionary genetics of the genes underlying phenotypic associations for loblolly pine (*Pinus taeda*, Pinaceae). Genetics 195 (4), 1353–1372. https://doi:10.1534/genetics.113.157198.

Elliot, W.J., Page-Dumroese, D., Robichaud, P.R., 2019. The effects of forest management on erosion and soil productivity*. USDA Forest Service RMRS, 195–208. https://doi:10.1201/9780203739266-12.

Enache, A., Kühmaier, M., Visser, R., and Stampfer, K., 2016. Forestry operations in the European mountains: a study of current practices and efficiency gaps. Scandinavian Journal of Forest Researc*h* 31, 412–427. https://doi:10.1080/02827581.2015.1130849.

Fatokun, C., Girma, G., Abberton, M., Gedil, M., Unachukwu, N., Oyatomi, O., Yusuf, M., Rabbi I., Boukar O., 2018. Genetic diversity and population structure of a mini-core subset from the world cowpea (*Vigna unguiculata* (L.) Walp.) germplasm collection. Scientific Reports 8, 1–10. https://doi:10.1038/s41598-018-34555-9.

Firn, J., Erskine, P. D., Lamb, D., 2007. Woody species diversity influences productivity and soil nutrient availability in tropical plantations. Oecologia 154, 521–533. https://doi:10.1007/s00442-007-0850-8.

Forester, B.R., Jones, M.R., Joost, S., Landguth, E.L., Lasky, J.R., 2016. Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. Molecular Ecology 25, 104–120. https://doi:10.1111/mec.13476.

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

François, O., Martins, H., Caye, K., Schoville, S.D., 2016. Controlling false discoveries in genome scans for selection. Molecular Ecology 25, 454–469. https://doi:10.1111/mec.13513.

Frichot, E., Schoville, S.D., Bouchard, G., Francois, O., 2013. LFMM version 1.2 - Reference Manual. 1–22.

Gehring, C.A., Sthultz, C.M., Flores-Rentería, L., Whipple, A.V., Whitham, T.G., 2017. Tree genetics defines fungal partner communities that may confer drought tolerance. PNAS 114 (42), 11169–11174. https://doi.org/10.1073/pnas.1704022114

George, J.P., Grabner, M., Karanitsch-Ackerl, S., Mayer, K., Weißenbacher, L., Schueler, S., 2017. Genetic variation, phenotypic stability, and repeatability of drought response in European larch throughout 50 years in a common garden experiment. Tree Physiology 37, 33–46. https://doi:10.1093/treephys/tpw085.

Gienapp P., Calus M.P.L., Laine V.N., Visser M.E. 2019. Genomic selection on breeding time in a wild bird population. Evol. Lett. 3, 142–151. https://doi:10.1002/evl3.103.

Gömöry, D., Comps, B., Paule, L., Von Wühlisch, G., 2013. Allozyme and phenotypic variation in beech (Fagus sylvatica L.): Are there any links? Plant Biosystems 147, 265–271. https://doi:10.1080/11263504.2013.763864.

Goudet J., Jombart T. 2015. Hierfstat: Estimation and tests of hierarchical F-statistics.

Gray, L. K., Hamann, A. 2011. Strategies for reforestation under uncertain future climates: Guidelines for Alberta, Canada. PLoS ONE 6 (8), e22977. https://doi:10.1371/journal.pone.0022977.

Guerrero, J., Andrello, M., Burgarella, C., Manel, S., 2018. Soil environment is a key driver of adaptation in Medicago truncatula: new insights from landscape genomics. New Phytologist 219, 378–390. https://doi:10.1111/nph.15171.

Hankin, L.E., Higuera, P.E., Davis, K.T., Dobrowski, S.Z., 2018. Accuracy of node and bud-scar counts for aging two dominant conifers in western North America. Forest Ecology and Management 427, 365–371. https://doi:10.1016/j.foreco.2018.06.001.

Harter, D.E.V., Nagy, L., Backhaus, S., Beierkuhnlein, C., Fussi, B., Huber, G., et al., 2015. A comparison of genetic diversity and phenotypic plasticity among european beech (Fagus sylvatica L.) populations from bulgaria and germany under drought and temperature manipulation. International Journal of Plant Sciences 176, 232–244.

https://doi:10.1086/679349.
Hoffmann, W. A., Poorter, H., 2002. Avoiding bias in calculations of relative growth rate. Annals of Botany 90, 37–42. https://doi:10.1093/aob/mcf140.
INEGI, Instituto Nacional de Estadística y Geografía (México)., 2015. Guía para la interpretación de cartografía : uso del suelo y vegetación : escala.1:250, 000 : serie V / Instituto Nacional de Estadística y Geografía.- México. https://www.inegi.org.mx/temas/usosuelo/.
Isabel, N., Holliday, J.A., Aitken, S.N., 2020. Forest genomics: Advancing climate adaptation, forest health, productivity, and conservation. Evolutionary Applications 13, 3–10. https://doi:10.1111/eva.12902.
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., González-Martínez S. C., 2015, Molecular Proxies for Climate Maladaptation in a Long-Lived Tree (*Pinus pinaster* Aiton, Pinaceae). Genetics 199 (3), 793-807. https://doi:10.1534/genetics.114.173252.
John, R., Dalling, J.W., Harms, K.E. et al. 2007. Soil nutrients influence spatial distributions of tropical tree species. PNAS 104 (3), 864-869. https://doi:10.1073/pnas.0604666104.
Kassambara, A., Mundt, F., 2020. Extract and visualize the results of multivariate data analyses. Version 1.0.7
Koskela, J., Vinceti, B., Dvorak, W., Bush, D., Dawson, I. K., Loo, J., et al., 2014. Utilization and transfer of forest genetic resources: A global review. Forest Ecology and Management 333, 22–34. https://doi:10.1016/j.foreco.2014.07.017.
Kozak, M., Bocianowski, J., Liersch, A., Tartanus, M., Bartkowiak-Broda, I., Piotto, F. A., et al., 2011. Genetic divergence is not the same as phenotypic divergence. Molecular Breeding 28, 277–280. https://doi:10.1007/s11032-011-9583-9.
Krasilnikov, P., Gutiérrez-Castorena, M.C., Ahrens, R.J., Cruz-Gaistardo, C.O., Sedov, S., Solleiro-Rebolledo, E., 2013. The soils of Mexico. Springer. Dordrecht, Netherlands.
Kubota, S., Iwasaki, T., Hanada, K., Nagano, A.J., Fujiyama, A., Toyoda, A., Sugano, S., Suzuki, Y., Hikosaka, K., Ita, M., Morinaga, S.I., 2015. A genome scan for genes underlying microgeographic-scale local adaptation in a wild Arabidopsis species. PLoS Genetics 11, 1–26. https://doi:10.1371/journal.pgen.1005361.
Lafleur, B., Labrecque, M., Arnold, A.A., Bélanger, N., 2015. Organic carbon accumulation in topsoil following afforestation with willow: Emphasis on leaf litter decomposition and soil organic matter quality. Forests 6, 769–793. https://doi:10.3390/f6030769.
Lang, R., 1920. Verwitterung und Bodenbildung als Einfuehrung in die Bodenkunde. Schweizer- bart Science Publishers, Stuttgart.
Larkin, D., Bruland, G., Zedler, J.B., 2016. Heterogeneity theory and ecological restoration 271-300. https://doi:10.5822/978-1-61091-698-1_10. Foundations of Restoration Ecology: Second Edition. Edition: 2nd Chapter: 10
Lascoux, M., Glémin, S., Savolainen, O., 2016. Local Adaptation in Plants. eLS, 1–7. https://doi:10.1002/9780470015902.a0025270.
Lawson, S.S., Michler, C.H. 2014., Afforestation, restoration and regeneration - not all trees are created equal. Journal of Forestry Research 25, 3–20. https://doi:10.1007/s11676-014-0426-5.
Le, H.D., Smith, C., Herbohn, J., Harrison, S., 2012. More than just trees: Assessing reforestation success in tropical developing countries. Journal of Rural Studies 28, 5–19. https://doi:10.1016/j.jrurstud.2011.07.006.
Le Corre, V., Kremer, A., 2012. The genetic differentiation at quantitative trait loci under local adaptation. Molecular Ecology 21 (7), 1548–1566. https://doi.org/10.1111/j.1365-294X.2012.05479.x
Lebedev, V.G., Lebedeva, T.N., Chernodubov, A.I., Shestibratov, K.A., 2020. Genomic selection for forest tree improvement: Methods, achievements and perspectives. Forests 11, 1–36. https://doi:10.3390/f11111190.
Lee, E., Li, Q., Eu, S., Han, S.K., Im, S., 2020. Assessing the impacts of log extraction by typical small shovel logging system on soil physical and hydrological properties in the Republic of Korea. Heliyon 6, e03544. https://doi:10.1016/j.heliyon.2020.e03544.
Legendre P. Legendre L. 2012. Numerical Ecology. Volume 24. 3rd Edition.
Lenz, P.R.N., Nadeau, S., Mottet, M.J., Perron, M., Isabel, N., Beaulieu, J., et al., 2020. Multi-trait genomic selection for weevil resistance, growth, and wood quality in Norway spruce. Evolutionary Applications 13, 76–94. https://doi:10.1111/eva.12823.
Li, D., Niu, S., Luo, Y., 2012. Global patterns of the dynamics of soil carbon and nitrogen stocks following afforestation: A meta-analysis. New Phytologist 195, 172–181. https://doi:10.1111/j.1469-8137.2012.04150.x.
Li, Y., Wu, X., Chen, T., Wang, W., Liu, G., Zhang, W., et al., 2018. Plant phenotypic traits eventually shape its microbiota: a common garden test. Frontiers in Microbiology 9, 1–13. https://doi:10.3389/fmicb.2018.02479.
Lin, Y.T., Whitman, W.B., Coleman, D.C., Chiu, C.Y., 2018. Effects of reforestation on the structure and diversity of bacterial communities in subtropical low mountain forest soils. Frontiers in Microbiology 9, 1–10.

https://doi:10.3389/fmicb.2018.01968.
Liu, M., Sui, X., Hu, Y., Feng, F., 2019. Microbial community structure and the relationship with soil carbon and nitrogen in an original Korean pine forest of Changbai Mountain, China. BMC microbiology 19, 218. https://doi:10.1186/s12866-019-1584-6.
Liu, M., Wang, Z., Li, S., Lu X., Wang, X. Han X., 2017. Changes in specific leaf area of dominant plants in temperate grasslands along a 2500-km transect in northern China. Sci Rep 7, 10780. https://doi.org/10.1038/s41598-017-11133-z
Lojewski, N.R., Fischer, D.G., Bailey, J.K., Schweitzer, J.A., Whitham, T.G., Hart, S.C., 2009. Genetic basis of aboveground productivity in two native *Populus* species and their hybrids. Tree Physiology 29, 1133–1142. https://doi:10.1093/treephys/tpp046.
Loo, J., Souvannavong, O., Dawson, I.K., 2014. Seeing the trees as well as the forest: The importance of managing forest genetic resources. *Forest Ecology and Management* 333, 1–8. https://doi:10.1016/j.foreco.2014.08.014.
Lotterhos, K.E., Whitlock, M.C., 2015. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. Molecular Ecology 24, 1031–1046. https://doi:10.1111/mec.13100.
Lu, P., Parker, W.C., Colombo, S.J., Man, R., 2016. Restructuring tree provenance test data to conform to reciprocal transplant experiments for detecting local adaptation. Journal of Applied Ecology 53, 1088–1097. https://doi:10.1111/1365-2664.12647.
Lu, X., Mo, J., Gilliam, F.S., Fang, H., Zhu, F., Fang, Y., et al., 2012. Nitrogen addition shapes soil phosphorus availability in two reforested tropical forests in Southern China. Biotropica 44, 302–311. https://doi:10.1111/j.1744-7429.2011.00831.x.
Ma, X., Mahecha, M.D., Migliavacca, M., van der Plas, F., Benavides, R., Ratcliffe, S., Kattge, J., Richter, R., Musavi T., Baeten, L., Barnoaiea T., Bohn F.J., Bouriaud, O., Bussotti, F., Coppi, A., Domisch, T., Huth, A., Jaroszewicz, B., Joswig, J., Pabon-Moreno J., Papale, D., Selvi, F., Laurin, G.V., Valladares, F., Reichstein, M., Wirth, C., 2019. "Inferring Plant Functional Diversity from Space : The Potential of Sentinel-2.". Remote Sensing of Environment 233, 111368.
Madritch, M.D., Lindroth, R.L., 2011. Soil microbial communities adapt to genetic variation in leaf litter inputs. Oikos 120, 1696–1704. https://doi:10.1111/j.1600-0706.2011.19195.x.
Mahony C.R., MacLachlan I.R., Lind, B.M., Yoder, J.B., Wang, T., Aitken, S.N., 2019. Evaluating genomic data for management of local adaptation in a changing climate: A lodgepole pine case study. Evolutionary Applications 13 (1), 116-131. https://doi.org/10.1111/eva.12871
Méndez-González, I.D., Jardón-Barbolla, L., Jaramillo-Correa, J.P., 2017. Differential landscape effects on the fine-scale genetic structure of populations of a montane conifer from central Mexico. Tree Genetics & Genomes 1, 13–30. https://doi.org/10.1007/s11295-017-1112-5
Mckown, A.D., Guy, R.D., Klápště, J., Geraldes, A., Friedmann, M., Cronk, Q.C.B., El-Kassaby Y.A., 2014. Geographical and environmental gradients shape phenotypic trait variation and genetic structure in *Populus trichocarpa*. New Phytologist 201, 1263–1276. https://doi:10.1111/nph.12601.
Minamikawa, M. F., Takada, N., Terakami, S., Saito, T., Onogi, A., Kajiya-Kanegae, H., Hayashi, T., Yamamoto, T., Iwata, H., 2018. Genome-wide association study and genomic prediction using parental and breeding populations of Japanese pear (*Pyrus pyrifolia*, Nakai). Scientific Reports 8, 11994. https://doi:10.1038/s41598-018-30154-w.
Mitchell-Olds, T., Willis, J.H., Goldstein, D.B., 2007. Which evolutionary processes influence natural genetic variation for phenotypic traits? Nature Reviews Genetics 8, 845–856. https://doi:10.1038/nrg2207.
Mosca, E., Cruz, F., Gómez-Garrido, J., Bianco, L., Rellstab, C., Brodbeck, S., … Neale, D. B., 2019. A reference genome sequence for the European Silver Fir (Abies alba Mill.): A community-generated genomic resource . G3; Genes Genomes Genetics, 9 (7), 2039–2049. https://doi.org/10.1534/g3.119.400083
Muthoo, M., 2002. Mountain environment & development. Unasylva 53, 26–35.
Nadeau, S., Meirmans, P.G., Aitken, S.N., Ritland, K., 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, 8649–8664. https://doi:10.1002/ece3.2550.
Oksanen, A.J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., et al., 2019. Package 'vegan'.
Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A., Ketudat Cairns, J.R., 2006. Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 β-glucosidase. BMC Plant Biology 6, 1–19. https://doi:10.1186/1471-2229-6-33.
Ortiz-Bibian, M.A., Blanco-García, A., Lindig-Cisneros, R.A., Gómez-Romero, M., Castellanos-Acuña, D., Herrerías-Diego, Y., Sánchez-Várgas, N.M., Sáenz-Romero, C., 2017. Genetic variation in Abies religiosa for quantitative traits and delineation of elevational and climatic zoning for maintaining monarch butterfly overwintering sites in mexico, considering climatic change. Silvae Genetica. 66, 14–23. https://doi:10.1515/sg-2017-0003.

Partida-Martínez, L.P.P., Heil, M., 2011. The microbe-free plant: fact or artifact? Front. Plant Sci. 2:100. https://doi: 10.3389/fpls.2011.00100.
Peres-neto, A.P.R., Legendre, P., Dray, S., Borcard, D., 2006. Variation partitioning of species data matrices. Ecology 87, 2614–2625.
Pluess, A.R., Frank, A., Heiri, C., Lalagüe, H., Vendramin, G.G., Oddou-Muratorio, S., 2016. Genome-environment association study suggests local adaptation to climate at the regional scale in Fagus sylvatica. New Phytologist 210, 589–601. https://doi:10.1111/nph.13809.
Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., et al. 2012. Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genome 5, 103–113. https://doi:10.3835/plantgenome2012.06.0006.
Poltorak, B.J., Labelle, E.R., Jaeger, D., 2018. Soil displacement during ground-based mechanized forest operations using mixed-wood brush mats. Soil and Tillage Researc*h* 179, 96–104. https://doi:10.1016/j.still.2018.02.005.
Poorter, H. 2002. Plant Growth and Carbon Economy. Encyclopedia of Life Sciences. https://doi:10.1038/npg.els.0003200.
Pregitzer, C.C., Bailey, J.K., Schweitzer, J.A., 2013. Genetic by environment interactions affect plant-soil linkages. Ecology and Evolution 3, 2322–2333. https://doi:10.1002/ece3.618.
Purahong, W., Durka, W., Fischer, M., Dommert, S., Schöps, R., Buscot, F., Wubet, T., 2016. Tree species, tree genotypes and tree genotypic diversity levels affect microbe-mediated soil ecosystem functions in a subtropical forest. Scientific Reports, 6 (36672), 1–11. https://doi.org/10.1038/srep36672
Purcell, S., 2010. Plink-Doc-1.07. Book, 1–293. Available at: papers://55069ee6-504c-4f60-bfa9-053c4dcabb39/Paper/p904.
R Core Team, 2016. R: A language and environment for statistical computing.
Ramírez-Valiente, J.A., Valladares, F., Aranda, I., 2014. Exploring the impact of neutral evolution on intrapopulation genetic differentiation in functional traits in a long-lived plant. Tree Genetic & Genomes 10: 1181–1190
Raven, J.A., Andrews, M., 2010. Evolution of tree nutrition. Tree Physiology 30, 1050–1071. https://doi:10.1093/treephys/tpq056.
Rellstab, C., Fischer, M.C., Zoller, S., Graf, R., Tedder, A., Shimizu, K.K., et al., 2017. Local adaptation (mostly) remains local: Reassessing environmental associations of climate-related candidate SNPs in Arabidopsis halleri. Heredity 118, 193–201. https://doi:10.1038/hdy.2016.82.
Rellstab, C., Gugerli, F., Eckert, A.J., Hancock, A.M., Holderegger, R., 2015. A practical guide to environmental association analysis in landscape genomics. Molecular Ecology 24, 4348–4370. https://doi:10.1111/mec.13322.
Ren, C., Sun, P., Kang, D., Zhao, F., Feng, Y., Ren, G., et al., 2016. Responsiveness of soil nitrogen fractions and bacterial communities to afforestation in the Loess Hilly Region (LHR) of China. Scientific Reports 6, 1–11. https://doi:10.1038/srep28469.
Riordan E.C., Gugger P.F., Ortego J., Smith C., Gaddis K., Thompson P., et al., 2016. Association of genetic and phenotypic variability with geography and climate in three southern California oaks. Am. J. Bot. 103, 73–85. https://doi:10.3732/ajb.1500135.
Roa-Fuentes, L.L., Martínez-Garza C., Etchevers J., Campo J., 2015. Recovery of soil C and N in a tropical pasture: passive and active restoration. Land Degradation & Development 26 (3), 201-210
Robertson, G., Wedin, D., Groffman, P., Blair, J., Holland, E., Nadelhoffer, K., Harris, D.F., 1999. Soil carbon and nitrogen availability: Nitrogen mineralization, nitrification, and soil respiration potentials. New York: Oxford University Press Oxford University Press.
Rzedowski, J., Fryxell, P.A., 2006. Vegetación de México (índice). Taxon 31, 793. https://doi:10.2307/1219727.
Sáenz-Romero, C., Lindig-Cisneros, R.A., Joyce, D.G., Beaulieu, J., Bradley, J.S.C., Jaquish, B. C., 2016. Assisted migration of forest populations for adapting trees to climate change. Revista Chapingo, Serie Ciencias Forestales y del Ambiente 22, 303–323. https://doi:10.5154/r.rchscfa.2014.10.052.
Sáenz-Romero, C., Rehfeldt, G.E., Soto-Correa, J.C., Aguilar-Aguilar, S., Zamarripa-Morales, V., López-Upton, J., 2012. Altitudinal genetic variation among Pinus pseudostrobus populations from michoacán, México. Two location shadehouse test results. Revista Fitotecnia Mexicana 35, 111–120.
Sáez-Plaza, P., Michałowski, T., Navas, M.J., Asuero, A.G., Wybraniec, S., 2013. An Overview of the Kjeldahl Method of nitrogen determination. Part I. Early history, chemistry of the procedure, and titrimetric finish. Critical Reviews in Analytical Chemistry 43, 178–223. https://doi:10.1080/10408347.2012.751786.
Sauer, T.J., James, D.E., Cambardella, C.A., Hernandez-Ramirez, G., 2012. Soil properties following reforestation or afforestation of marginal cropland. Plant and Soil 360, 375–390. https://doi:10.1007/s11104-012-1258-8.
Savolainen, O., Lascoux, M., Merilä, J., 2013. Ecological genomics of local adaptation. Nature Reviews Genetics 14, 807–820. https://doi:10.1038/nrg3522.
Savolainen, O., Pyhäjärvi, T., Knürr, T., 2007. Gene flow and local adaptation in trees. Annual Review of Ecology, Evolution, and Systematics 38:595–619. https://doi.org/10.1146/annurev.ecolsys.38.091206.095646

Schweitzer, J. A., Fischer, D.G., Rehill, B.J., Wooley, S.C., Woolbright, S.A., Lindroth, R.L., et al., 2011. Forest gene diversity is correlated with the composition and function of soil microbial communities. Population Ecology 53, 35–46. https://doi:10.1007/s10144-010-0252-3.

Scotti, I., González-Martínez, S.C. , Budde K.B. 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, 69–75. https://doi:10.1007/s13595-015-0471-z

Sebastian-Azcona J., Hacke U., Hamann A., 2020. Xylem anomalies as indicators of maladaptation to climate in forest trees: Implications for assisted migration. Frontiers in Plant Science. 11 (208). https://doi:10.3389/fpls.2020.00208.

Sheoran, V., Sheoran, A.S., Poonia, P., 2016. Factors Affecting Phytoextraction: A review. Pedosphere 26 (2), 148-166. https://doi.org/10.1016/S1002-0160(15)60032-7.

Silfver, T., Paaso, U., Rasehorn, M., Rousi, M., Mikola, J., 2015. Genotype × herbivore effect on leaf litter decomposition in Betula pendula saplings: Ecological and evolutionary consequences and the role of secondary metabolites. PLoS ONE 10, 1–15. https://doi:10.1371/journal.pone.0116806.

Silva, L.C.R., Lambers, H., 2020. Soil-plant-atmosphere interactions: structure, function, and predictive scaling for climate change mitigation. Plant Soil. https://doi.org/10.1007/s11104-020-04427-1.

Skotte, L., Korneliussen, T.S., Albrechtsen, A., 2013. Estimating individual admixture proportions from next generation sequencing data. Genetics 195, 693–702. https://doi:10.1534/genetics.113.154138.

Soltys-Kalina, D., Plich, J., Strzelczyk-Żyta, D., Śliwka, J., Marczewski, W., 2016. The effect of drought stress on the leaf relative water content and tuber yield of a half-sib family of 'Katahdin'-derived potato cultivars. Breeding Science 66, 328–331. https://doi:10.1270/jsbbs.66.328.

Sork, V.L., 2018. Genomic studies of local adaptation in natural plant populations. Journal of Heredity. 109, 3–15. https://doi:10.1093/jhered/esx091.

Sork, V.L., Aitken, S.N., Dyer, R.J., Eckert, A.J., Legendre, P., 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, 901–911. https://doi:10.1007/s11295-013-0596-x.

Spokevicius, A.V., Tibbits, J., Rigault, P., Nolin, M.A., Müller, C., Merchant, A., 2017. Medium term water deficit elicits distinct transcriptome responses in Eucalyptus species of contrasting environmental origin. BMC Genomics 18 (1), 284.

Stanton-geddes, J., Yoder, J.B., Briskine, R., Young, N.D., Tiffin, P., 2013. Estimating heritability using genomic data. Methods in Ecology and Evolution 4, 1151–1158. https://doi:10.1111/2041-210X.12129.

Stouffer, S.A., Suchman, E.A., De Vinney, L.C., Star, S.A., Williams, R.M. Jr., 1949. The American Soldier, Vol.1: Adjustment during Army Life. Princeton University Press, Princeton.

Storey, J. D., Tibshirani, R., 2003. Statistical significance for genomewide studies. PNAS. 100 (16), 9440-9445. https://doi:10.1073/pnas.1530509100.

Talbot, B., Chen, T.W., Zimmerman, S., Joost, S., Eckert, A.J., Crow, T.M., et al., 2017. Combining genotype, phenotype, and environment to infer potential candidate genes. Journal of Heredity. 108, 207–216. https://doi:10.1093/jhered/esw077.

Telfer, E., Graham, N., Stanbra, L., Manley, T., Wilcox, P., 2013. Extraction of high purity genomic DNA from pine for use in a high-throughput genotyping platform. New Zealand Journal of Forestry Science 43:3.

ter Braak, C.J.F., Verdonschot, P.F.M., 1995. Canonical correspondence analysis and related multivariate methods in aquatic ecology. Aquatic Sciences 57, 255–289. https://doi:10.1007/BF00877430.

Tiffin, P., Ross-Ibarra, J., 2014. Advances and limits of using population genetics to understand local adaptation. Trends in Ecology and Evolution 29 (12), 673–680. https://doi.org/10.1016/j.tree.2014.10.004

Tolkamp, G.W., Priadjati, A., Effendi, R., 1999. Towards an ecology-based strategy for the reforestation of imperata cylindrica grasslands in. The balance between biodiversity conservation and sustainable use of tropical rain forest*s* 99–116.

Turner, S.D., 2018. qqman : an R package for visualizing GWAS results using Q-Q and manhattan plots. 3, 2–3. https://doi:10.21105/joss.00731.

Ueno, S., Nakamura, Y., Kobayashi, M., Terashima, S., Ishizuka, W., Uchiyama, K., … Goto, S., 2018. TodoFirGene: Developing transcriptome resources for genetic analysis of Abies sachalinensis. Plant and Cell Physiology 59 (6), 1276–1284. https://doi.org/10.1093/pcp/pcy058

Urrestarazu, J., Muranty, H., Denancé, C., Leforestier, D., Ravon, E., Guyader, A., et al., 2017. Genome-wide association mapping of flowering and ripening periods in apple. Frontiers in Plant Science 8, 1–19. https://doi:10.3389/fpls.2017.01923.

Urza, A.K., Sibold, J.S., 2013. Nondestructive aging of postfire seedlings for four conifer species in northwestern montana. Western Journal of Applied Forestry 28, 22–29. https://doi:10.5849/wjaf.11-014.

VanRaden P.M., 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91, 4414-4423. https://doi.org/10.3168/jds.2007-0980.
Villellas, J., Berjano, R., Terrab, A., García, M.B., 2014. Divergence between phenotypic and genetic variation within populations of a common herb across Europe. Ecosphere 5, 1–14. https://doi:10.1890/ES13-00291.1.
Vizcaíno-Palomar, N., Revuelta-Eugercios, B., Zavala, M.A., Alía, R., Gonzalez-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. https://doi:10.1371/journal.pone.0109132.
Wachowiak, W., Perry, A., Donnelly, K., Cavers, S., 2018. Early phenology and growth trait variation in closely related european pine species. Ecology and Evolution 8, 655–666. https://doi:10.1002/ece3.3690.
Wang, M.Z., Li, H.L., Li, J.M., Yu, F.H., 2020. Correlations between genetic, epigenetic and phenotypic variation of an introduced clonal herb. Heredity 124, 146–155. https://doi:10.1038/s41437-019-0261-8.
Wardle D.A., Walker L.R., Bardgett R.D., 2004. Ecosystem properties and forest decline in contrasting long-term chronosequences. Science 305, 509-513.
Weir B.S., Cockerham C.C., 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1358–70.
Wickham, H., 2019. Package ‘ tidyverse .’ 1–5.
Williamson, G.B., Wiemann, M.C., 2010. Measuring wood specific gravity... correctly. American Journal of Botany 97, 519–524. https://doi:10.3732/ajb.0900243.
Wills, J., Herbohn, J., Moreno, M.O.M., Avela, M.S., Firn, J., 2017. Next-generation tropical forests: reforestation type affects recruitment of species and functional diversity in a human-dominated landscape. Journal of Applied Ecology 54, 772–783. https://doi:10.1111/1365-2664.12770.
Zhang, M., Suren, H., Holliday, J. A. 2019. Phenotypic and genomic local adaptation across latitude and altitude in Populus trichocarpa. Genome biology and evolution 11, 2256–2272. https://doi:10.1093/gbe/evz151.
Zhou, G., Zhang, J., Zhang, C., Feng, Y., Chen, L., Yu, Z., et al., 2016. Effects of changes in straw chemical properties and alkaline soils on bacterial communities engaged in straw decomposition at different temperatures. Scientific Reports 6. https://doi:10.1038/srep22186.
Zianis, D., Muukkonen, P., Mäkipää, R., Mencuccini, M., 2005. Biomass and stem volume equations for tree species in europe. Silva Fennica Monographs. 4. pp. 63.

TABLES

Table 1. Mean and standard error (SE) for nine soil traits measured in a sacred fir (*Abies religiosa*) trial site of in central Mexico, followed by *t-test* results (*t- value* and significance of *P*-value) performed between two groups of individuals; natural regeneration (NR), and introduced sapling (RF).

Soil characteristic	***Natural regeneration***			***Reforestation***			***t-test***	***P*-value**
	Mean		**SE**	**Mean**		**SE**		
pH (H_2O)	5.8	±	0.1	6.3	±	0.04	-5.14	***
Organic carbon (C, mg/g)	187.6	±	7.8	159.2	±	6.1	3.76	***
Total Nitrogen (TN, mg N /g)	18.5	±	1.1	14.3	±	0.5	3.60	***
Organic carbon/Total nitrogen ratio (C:TN)	10.6	±	0.3	11.1	±	0.4	-0.10	Ns
Nitrate (NO_3^-, µg/g)	8.29	±	1.27	9.97	±	0.70	-1.28	Ns
Ammonium (NH_4^+, µg/g)	27.9	±	2.2	40.9	±	2.1	-4.29	***
Nitrate/Ammonium ratio (NO_3^-:NH_4^+)	0.28	±	0.02	0.27	±	0.03	-0.07	Ns
Total Phosphorus (TP, mg P/g)	2.65	±	0.17	2.59	±	0.18	0.20	Ns
Total Nitrogen /Total Phosphorus ratio (TN:TP)	7.43	±	0.38	6.44	±	0.43	1.96	Ns

Ns, non-significant; *, *P*-value < 0.05; **, *P*-value < 0.01; ***, *P*-value < 0.001

Table 2. Results (variance constrained and unconstrained) of Redundancy Analyses (RDA) for accounting for the variation at three sets of quantitative traits (models) with 1,749 SNPs. $Model_{G1}$, growth traits only; $Model_{G2}$, physiological traits only; $Model_{G3}$, all traits. Fisher's *F*, number of degrees of freedom (Df), *P*-value and adjusted coefficient of determination R^2 (R^2_{adj}) are shown also for each model at the bottom of the table.

	$Model_{G1}$	$Model_{G2}$	$Model_{G3}$
Variance-Constrained (%)	10.91	6.86	16.97
Variance-Unconstrained (%)	88.09	93.14	83.03
F-test	1.07	1.10	1.05
Df	6	4	10
P	0.08^{Ns}	0.06^{Ns}	0.12^{Ns}
R^2_{adj}	0.0059	0.0065	0.0051

Ns, non-significant; *, *P*-value < 0.05

Table 3. Results (% variance explained) for partial canonical correlation analyses (pCCA) between genomic and edaphic variation and three sets of phenotypic traits (models) in a trial Sacred fir site of in Central Mexico. $Model_{G\times E1}$, growth traits; $Model_{G\times E2}$, physiological traits; $Model_{G\times E3}$, all traits.

Source of Explained Variance	**Contribution to total variance (%)**		
	$Model_{G\times E1}$	$Model_{G\times E2}$	$Model_{G\times E3}$
Effect genetical	10.02	8.28	9.8
Effect environmental	3.8	3.91	4.9
Effect genotype × environment	5.63	3.79	6.88
Total explained	19.45	15.98	21.58
Residual	80.55	84.02	78.42

Table 4. Results of the cross-validation of the model to predict the genomic and phenotypic variance calculated using predictive soil traits. The statistical measures are: To quantify the performance of the model obtained from the calibration data set, we used the calculated coefficient of determination (R^2_{Cal}), while from the validation data sets we obtained the determination coefficient (R^2_{CV}); and normalized mean square error ($nRMSE_{CV}$, RMSE divided by the sample mean), and in both cases it is indicated by the CV subscript.

Method	Model	Statistical measures	RDA1	RDA2	RDA3
Across-groups	NR → RF	R^2_{Cal}	0.382	0.391	0.335
		R^2_{CV}	0.085	0.137	0.069
		$nRMSE_{CV}$	2.32	13.34	4.38
Resampling	50 %'s sample	R^2_{Cal}	0.453	0.362	0.406
		R^2_{CV}	0.121	0.078	0.076
		$nRMSE_{CV}$	16.3	28.01	18.58
	75%'s sample	R^2_{Cal}	0.383	0.297	0.371
		R^2_{CV}	0.224	0.085	0.114
		$nRMSE_{CV}$	20.94	24.51	23.68
Compound	NR + $(0.5*RF_{TRN})$ → $0.5*RF_{TST}$	R^2_{Cal}	0.371	0.351	0.321
		R^2_{CV}	0.128	0.163	0.122
		$nRMSE_{CV}$	12.63	20.608	9.729

SUPPLEMENTARY MATERIAL

Table S1. Geographical coordinates for the 128 Sacred fir (*Abies religiosa*) individuals surveyed in this study.

Origin of Individual					
Origin	**Population**	**No. of Individual**	**Latitude**	**Longitude**	**Genetic and soil samples**
Reforestation	RF	51	19.2577	-99.9569	✓
Reforestation	RF	52	19.257	-99.9566	✗
Reforestation	RF	53	19.2572	-99.9565	✗
Reforestation	RF	54	19.257	-99.9566	✓
Reforestation	RF	55	19.257	-99.9565	✗
Reforestation	RF	56	19.2569	-99.9565	✗
Reforestation	RF	57	19.2567	-99.9566	✓
Reforestation	RF	58	19.2568	-99.9564	✗
Reforestation	RF	59	19.2568	-99.9566	✓
Reforestation	RF	60	19.2567	-99.9565	✗
Reforestation	RF	61	19.2566	-99.9566	✓
Reforestation	RF	62	19.2565	-99.9564	✗
Reforestation	RF	63	19.2564	-99.9564	✗
Reforestation	RF	64	19.2563	-99.9563	✓
Reforestation	RF	65	19.2562	-99.956	✗
Reforestation	RF	66	19.2563	-99.9561	✓
Reforestation	RF	67	19.2563	-99.956	✗
Reforestation	RF	68	19.2563	-99.9554	✗
Reforestation	RF	69	19.2563	-99.9561	✗
Reforestation	RF	70	19.2565	-99.9561	✗
Reforestation	RF	71	19.2566	-99.9562	✗
Reforestation	RF	72	19.2571	-99.9563	✓
Reforestation	RF	73	19.2575	-99.9568	✗
Reforestation	RF	74	19.2579	-99.957	✓
Reforestation	RF	75	19.2581	-99.9573	✗
Reforestation	RF	76	19.2576	-99.9566	✗
Reforestation	RF	77	19.2577	-99.9565	✓
Reforestation	RF	78	19.2577	-99.9566	✗
Reforestation	RF	79	19.258	-99.9566	✗
Reforestation	RF	80	19.2576	-99.9567	✓
Reforestation	RF	81	19.2576	-99.9566	✓
Reforestation	RF	82	19.2575	-99.9564	✗
Reforestation	RF	83	19.2576	-99.9564	✓
Reforestation	RF	84	19.2575	-99.9566	✗
Reforestation	RF	85	19.2574	-99.9565	✗
Reforestation	RF	86	19.2574	-99.9567	✓
Reforestation	RF	87	19.2571	-99.9566	✗
Reforestation	RF	88	19.2573	-99.9566	✗
Reforestation	RF	89	19.2568	-99.9563	✓

Reforestation	RF	90	19.256	-99.9561	✗
Reforestation	RF	91	19.2567	-99.9565	✓
Reforestation	RF	92	19.2567	-99.9563	✗
Reforestation	RF	93	19.2568	-99.9567	✓
Reforestation	RF	94	19.2567	-99.9563	✓
Reforestation	RF	95	19.2567	-99.9563	✗
Reforestation	RF	96	19.2568	-99.9564	✓
Reforestation	RF	97	19.2568	-99.9565	✗
Reforestation	RF	98	19.2569	-99.9565	✗
Reforestation	RF	99	19.2569	-99.9566	✗
Reforestation	RF	100	19.2574	-99.957	✓
Reforestation	RF	101	19.2581	-99.9572	✗
Reforestation	RF	102	19.2583	-99.9577	✓
Reforestation	RF	103	19.2583	-99.9575	✓
Reforestation	RF	104	19.2583	-99.9575	✓
Reforestation	RF	105	19.2583	-99.9576	✓
Reforestation	RF	106	19.2582	-99.9575	✗
Reforestation	RF	107	-	-	✗
Reforestation	RF	108	19.2575	-99.9564	✓
Reforestation	RF	109	19.2563	-99.9565	✗
Reforestation	RF	110	19.2566	-99.9562	✗
Reforestation	RF	111	19.2568	-99.9559	✓
Reforestation	RF	112	19.2567	-99.9558	✗
Reforestation	RF	113	19.2567	-99.9559	✗
Reforestation	RF	114	19.2561	-99.9564	✗
Reforestation	RF	115	19.2564	-99.9564	✗
Reforestation	RF	116	19.257	-99.9567	✓
Reforestation	RF	117	19.257	-99.9567	✗
Reforestation	RF	118	19.2572	-99.9556	✗
Reforestation	RF	119	19.2576	-99.9566	✗
Reforestation	RF	120	19.2578	-99.9565	✗
Reforestation	RF	121	19.2571	-99.9563	✓
Reforestation	RF	122	19.258	-99.9566	✓
Reforestation	RF	123	19.2581	-99.9565	✓
Reforestation	RF	124	19.2578	-99.9562	✓
Reforestation	RF	125	-	-	✗
Reforestation	RF	126	19.2572	-99.9565	✓
Reforestation	RF	127	-	-	✗
Natural regeneration	NR	1	19.2576	-99.9567	✓
Natural regeneration	NR	2	19.2572	-99.9564	✗
Natural regeneration	NR	3	19.2571	-99.9564	✓
Natural regeneration	NR	4	19.2571	-99.9568	✓
Natural regeneration	NR	5	19.2574	-99.9569	✓
Natural regeneration	NR	6	19.2573	-99.9568	✓
Natural regeneration	NR	7	19.2577	-99.9571	✗

Natural regeneration	NR	8	19.2579	-99.9573	✓
Natural regeneration	NR	9	19.2574	-99.9567	✓
Natural regeneration	NR	10	19.2565	-99.9563	✗
Natural regeneration	NR	11	19.2566	-99.9564	✓
Natural regeneration	NR	12	19.2566	-99.9563	✓
Natural regeneration	NR	13	19.2561	-99.9558	✓
Natural regeneration	NR	14	19.2575	-99.9564	✗
Natural regeneration	NR	15	19.258	-99.9571	✓
Natural regeneration	NR	16	19.2561	-99.9558	✓
Natural regeneration	NR	17	19.2567	-99.9564	✓
Natural regeneration	NR	18	-	-	✗
Natural regeneration	NR	19	19.2563	-99.9559	✓
Natural regeneration	NR	20	19.2575	-99.957	✓
Natural regeneration	NR	21	-	-	✗
Natural regeneration	NR	22	19.2573	-99.9567	✓
Natural regeneration	NR	23	19.2564	-99.9562	✓
Natural regeneration	NR	24	19.2581	-99.957	✓
Natural regeneration	NR	25	19.2575	-99.957	✓
Natural regeneration	NR	26	19.2577	-99.9571	✓
Natural regeneration	NR	27	19.2577	-99.9571	✗
Natural regeneration	NR	28	-	-	✗
Natural regeneration	NR	29	19.2574	-99.9569	✗
Natural regeneration	NR	30	19.2567	-99.9559	✓
Natural regeneration	NR	31	19.2564	-99.9558	✓
Natural regeneration	NR	32	19.2574	-99.9566	✓
Natural regeneration	NR	33	19.2573	-99.9567	✓
Natural regeneration	NR	34	19.2575	-99.9566	✓
Natural regeneration	NR	35	-	-	✗
Natural regeneration	NR	36	-	-	✗
Natural regeneration	NR	37	19.2574	-99.9564	✓
Natural regeneration	NR	38	19.2580	-99.9563	✓
Natural regeneration	NR	39	-	-	✗
Natural regeneration	NR	40	19.2554	-99.9555	✓
Natural regeneration	NR	41	-	-	✗
Natural regeneration	NR	42	-	-	✗
Natural regeneration	NR	43	-	-	✗
Natural regeneration	NR	44	-	-	✗
Natural regeneration	NR	45	19.2560	-99.9576	✓
Natural regeneration	NR	46	-	-	✗
Natural regeneration	NR	47	19.2557	-99.9554	✓
Natural regeneration	NR	48	19.2571	-99.9563	✓
Natural regeneration	NR	49	19.2566	-99.9559	✓
Natural regeneration	NR	50	19.2558	-99.9574	✓
Natural regeneration	NR	51	19.2575	-99.9574	✓

Table S2. Ten-fold cross-validation (CV) schemes on the three pre-established statistical models to the predictive power of the top soil variables on the RDA vectors in *Abies religiosa*. Predictive models were developed between a training and a validation dataset, which have different TRN/TST sizes.

Method	Model	TRN/TST
Across-groups	NR → RF	34/31
Resampling	50 %'s sample	33/32
	75%'s sample	48/17
Compound	NR + (0.5*RF_{TRN}) → 0.5*RF_{TST}	49/16

Table S3. Mean and standard error (SE) for the 21 phenotypic traits surveyed in a Sacred fir (*Abies religiosa*) trial site in central Mexico, followed by *t*-student results for the response of morphological and physiological phenotypes performed between two groups of individuals; natural regeneration (NR), and introduced saplings (RF).

Phenotypic trait	***Natural regeneration (NR)***			***Reforestation (RF)***			***t-test***	***P*-value**
	Mean		**SE**	**Mean**		**SE**		
Growth traits								
Total height (cm)	140.6	±	6.5	93.5	±	3.21	-6.46	***
Frond height (cm)	96.7	±	5.85	68.9	±	3.15	-4.19	***
Steam diameter (cm)	1.6	±	0.09	0.97	±	0.03	-6.76	***
Base diameter (cm)	1.95	±	0.1	1.3	±	0.04	-6.14	***
Above-ground biomass (kg)	582	±	65	198	±	16	-5.71	***
Crown radio (cm)	53.0	±	2.8	32.9	±	1.0	-7.46	***
Growth in 2017 (cm)	24.2	±	2.0	23.3	±	1.6	-0.36	Ns
Growth in 2016 (cm)	26.38	±	2.06	10.4	±	1.18	-6.73	***
Average anual growth (cm)	21.23	±	1.31	17.68	±	0.81	-2.29	*
Needle length (mm)	19.4	±	0.23	19.1	±	0.2	-0.76	Ns
Above-ground volumen (m^3)	0.38		0.05	0.085		0.007	-6.18	***
Wood density (g/cm^3)	0.54	±	0.02	0.41	±	0.01	-6.12	***
Physiological Traits								
Water potential (MPa)	-0.32	±	0.015	-0.47	±	0.02	-6.23	***
Relative water content (%)	74.6	±	1.9	65.7	±	1.4	-4.01	***
Specific leaf area (cm^2/g)	0.038	±	0.001	0.039	±	0.001	0.21	Ns
Relative growth rate ($g\ g^{-1}d^{-1}$)	1.32	±	0.15	1.54	±	0.15	1.16	Ns

Leaf N (mg N/g leaf)	19.3	±	2.6	22.9	±	2.05	-1.51	Ns
Leaf P (mg P/ g leaf)	1.28	±	0.1	1.81	±	0.09	-4.11	***
Leaf N:P ratio	15.4	±	2.9	12.5	±	0.9	1.34	Ns

*, *P*-value < 0.05; **, *P*-value < 0.01; ***, *P* -value < 0.001

Table S4. Eigenvalues, cumulative percent variation, and eigenvectors of the first six principal components (*PC*s) for ten plant traits surveyed in a sacred fir (*Abies religiosa*) trial site in central Mexico.

	Eigenvalues					
Components	**PC_1**	**PC_2**	**PC_3**	**PC_4**	**PC_5**	**PC_6**
Eigenvalue	2.97	1.91	0.97	0.79	0.7	0.56
Variance cumulative	33.09	54.36	69.16	80.01	86.25	90.93
Trait	**PC_1**	**PC_2**	**PC_3**	**PC_4**	**PC_5**	**PC_6**
Total height	0.71	0.49	0.03	-0.11	-0.21	-0.04
Steam diameter	0.59	0.53	0.14	-0.06	-0.16	-0.14
Crown radio	0.65	0.44	-0.07	-0.08	-0.05	-0.15
Growth in 2016	0.64	-0.24	-0.03	-0.46	0.6	0.42
Needle length	0.15	-0.31	0.91	-0.02	-0.19	0.02
Wood density	0.51	-0.55	0.09	0.41	-0.05	0.36
Water potential	0.6	-0.44	0.25	0.35	-0.31	-0.03
Relative water content	0.35	-0.64	-0.03	0.11	0.23	-0.41
Specific leaf area	-0.25	0.57	0.23	0.66	0.29	0.13
Leaf P	-0.42	0.45	0.3	0.56	-0.36	-0.1

Table S5. Results for a permutational multivariate analysis of variance (PERMANOVA) between the phenotypes of reforested and naturally regenerated individuals of sacred fir (*Abies religiosa*) surveyed in a trial site from central Mexico.

	Df	*Sum of Sqs*	R^2	*F*	*P*-value
Origin	1	100,650	0.952	2522	< 0.001
Residual	126	5,028	0.048		
Total	127	105,677	1.000		

Table S6. Cross-validation results for Admixture analyses of reforested and naturally regenerated individuals of sacred fir (*Abies religiosa*) surveyed in a trial site from central Mexico. *K*-values from 2 to 5 were assumed.

***K*- value**	**CV-error**
***K* =1**	0.509
***K* =2**	0.498
***K*=3**	0.521
***K*=4**	0.550
***K*=5**	0.596

Table S7. Results of a partial canonical correspondence analysis (pCCA) for testing for within-population genotype × environment interaction on three models of phenotypic traits evaluated in sacred fir (*Abies religiosa*) individuals from a trial site in central Mexico. *F*, Fisher's *F*; Df, degrees of freedom; *P*-value; R^2_{adj}, adjusted coefficient of determination.

	Model$_{G\times E1}$	**Model$_{G\times E2}$**	**Model$_{G\times E3}$**
Variance-Constrained	19.45	15.98	21.56
Variance-Unconstrained	80.55	84.02	78.44
F	2.35	1.98	2.69
Df	6	6	6
***P*-value**	0.02	0.08	0.002
R^2_{adj}	0.115	0.0840	0.1400

Table S8. Eigenvalues and cumulative variance for each axis of a variance partition analysis (pCCA) of phenotypic, genomic and edaphic traits evaluated in sacred fir (*Abies religiosa*) individuals from a trial site in central Mexico.

Components	**pCCA$_1$**	**pCCA$_2$**	**pCCA$_3$**	**pCCA$_4$**	**pCCA$_5$**
Eigenvalue	0.0042	0.0007	0.0002	0.0001	4.00E^{-05}
Variance cumulative	79.6	94.0	97.2	99.1	99.9

Table S9. Results for an analysis of variance partition analysis on the first five pCCA axes for accounting for the phenotypic variation of sacred fir (*Abies religiosa*) individuals from a trial site in central Mexico.

Components	*Df*	X^2	*F*	*P*-value
$pCCA_1$	1	0.0050	13.90	< 0.001
$pCCA_2$	1	0.0004	1.20	Ns
$pCCA_3$	1	0.0001	0.30	Ns
$pCCA_4$	1	0.0001	0.21	Ns
$pCCA_5$	1	0.0000	0.07	Ns
Residual	59	0.0206		

Ns: $P > 0.05$

Table S10. Results for a phenotypic variation model and statistical contribution of each variable (edaphic and genomic relationship eigenvalues) as estimated with Monte Carlo simulations for sacred fir (*Abies religiosa*) individuals from a trial site in central Mexico, based on a variance partition analysis (pCCA) followed by Fisher's *F* and significance level.

Predictors	*Df*	X^2	*F*	*P*-value
NH_4	1	0.0014	3.57	*
TN	1	0.0009	2.30	.
TP	1	0.0007	1.97	Ns
NO_3:NH_4 ratio	1	0.0004	1.31	Ns
GRE_1	1	0.0010	2.68	*
GRE_2	1	0.0012	3.48	*
Residual	58	0.0208		

., *P*-value < 0.1; *, *P*-value < 0.05; **

Table S11. L ist and putative annotation of contigs containing single nucleotide polymorphisms (SNPs) associated with soil properties in a sacred fir (*Abies religiosa*) trial site in central Mexico.

Soil properties	Associated SNPs	*q*-value	Gene Name	Funtion
Ammonium (NH_4)	*locus_634977*	1.30^{-2}	NA	NA
	locus_422812	1.99^{-2}	NA	NA
Total N (TN)	*locus_288495*	8.90^{-3}	-	NA
	locus_635539	8.95^{-3}	*Pentatricopeptide repeat (PPR) superfamily protein*	Protein amino acid binding
	locus_636923	8.95^{-3}	-	NA
	locus_176834	2.68^{-2}	*Glycosyl hydrolase 9B13*	Carbohydrate metabolic process
	locus_496243	2.70^{-2}	NA	NA
	locus_635920	3.87^{-2}	NA	NA
	locus_87407	3.91^{-2}	NA	NA
NO_3:NH_4 ratio	*locus_231253*	3.37^{-3}	NA	NA
	locus_628906	1.45^{-2}	NA	NA
	locus_291346	3.16^{-2}	NA	NA
	locus_542357	3.16^{-2}	NA	NA
	locus_567195	3.37^{-2}	NA	NA

SUPPLEMENTARY FIGURES

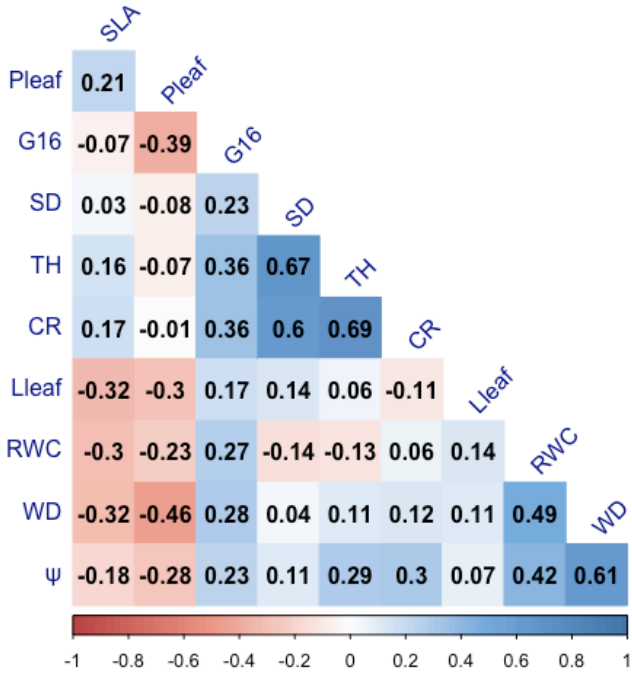

Fig. S1. Corplot showing statistically significant relationships between selected phenotypic traits in sacred fir (*Abies religiosa*). Numbers within boxes are Pearson's correlation coefficients. Red indicates a significantly positive correlation, blue a significantly negative correlation.

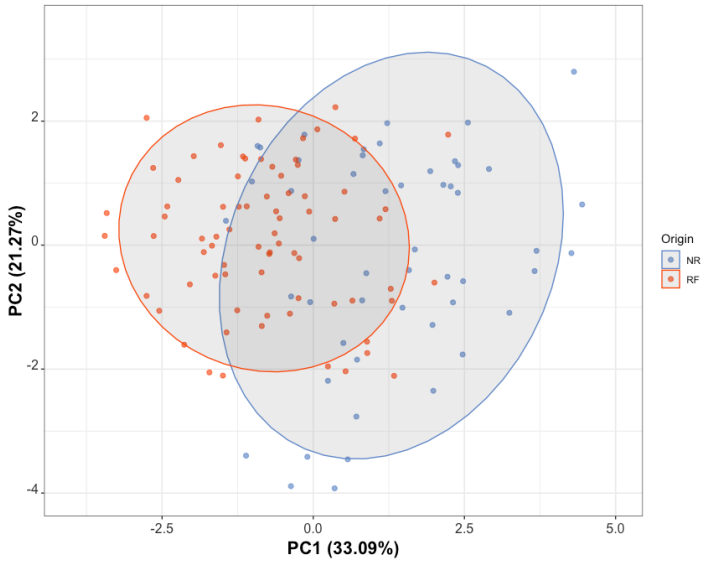

Fig. S2. PCA biplot for the phenotypic variation (growth and physiological traits) of 128 Sacred fir (*Abies religiosa*) individuals sampled in a trial test in Central Mexico. Individuals had two origins, natural regeneration (NR, blue dots) and an introduced plants from reforestation with foreign germplasm (RF, orange dots). The percentage of the phenotypic variation explained by each component is indicated at each axis.

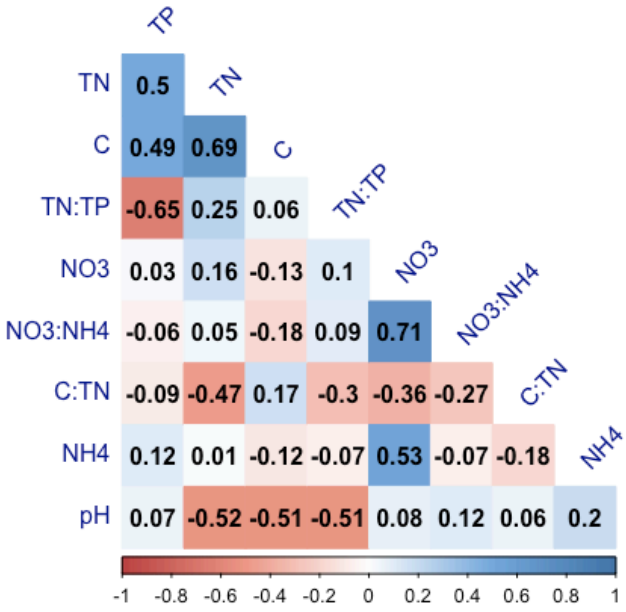

Fig S3: Corplot showing a statistically significant relationship between selected edaphic characteristics in a trial site of sacred fir (*Abies religiosa*) in central Mexico. Numbers within boxes are Pearson's correlation coefficients. Red indicates a significantly positive correlation, blue a significantly negative correlation.

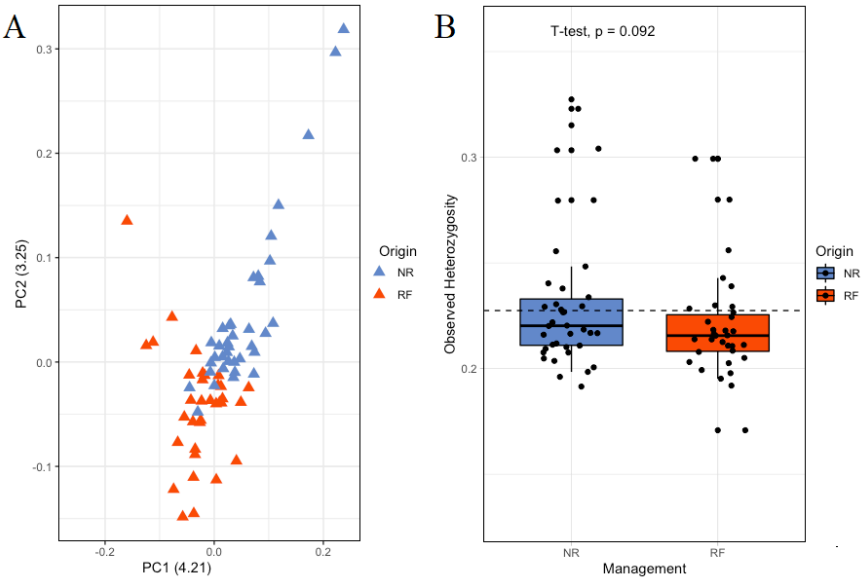

Fig S4 Genetic structure and diversity of studied sacred fir individuals based on 1749 SNPs. **A.** A principal component analysis (PCA) shows a weak genetic structure in a diversity panel of 65 individuals. The percentage of the variation explained by each component is indicated at each axis. **B.** Observed heterozygosity for each group of individuals. Red, introduced plants from reforestation (RF); blue, natural regeneration (NR).

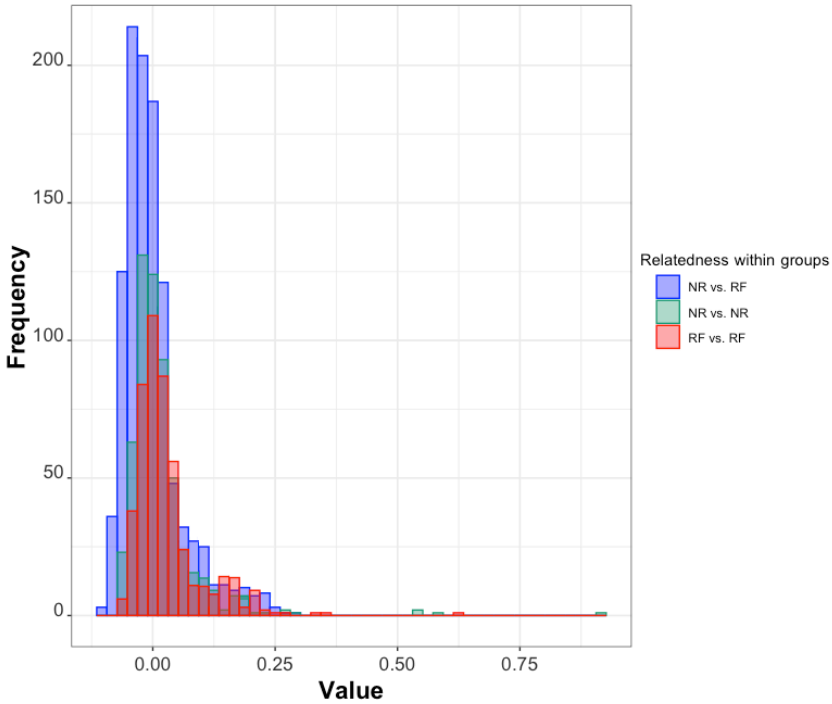

Fig. S5 Genetic relationship between the 65 sacred fir (*Abies religiosa*) individuals sampled in a trial site from central Mexico. The density plot of the genomic relationship matrix shows the absence of differentiated groups between samples.

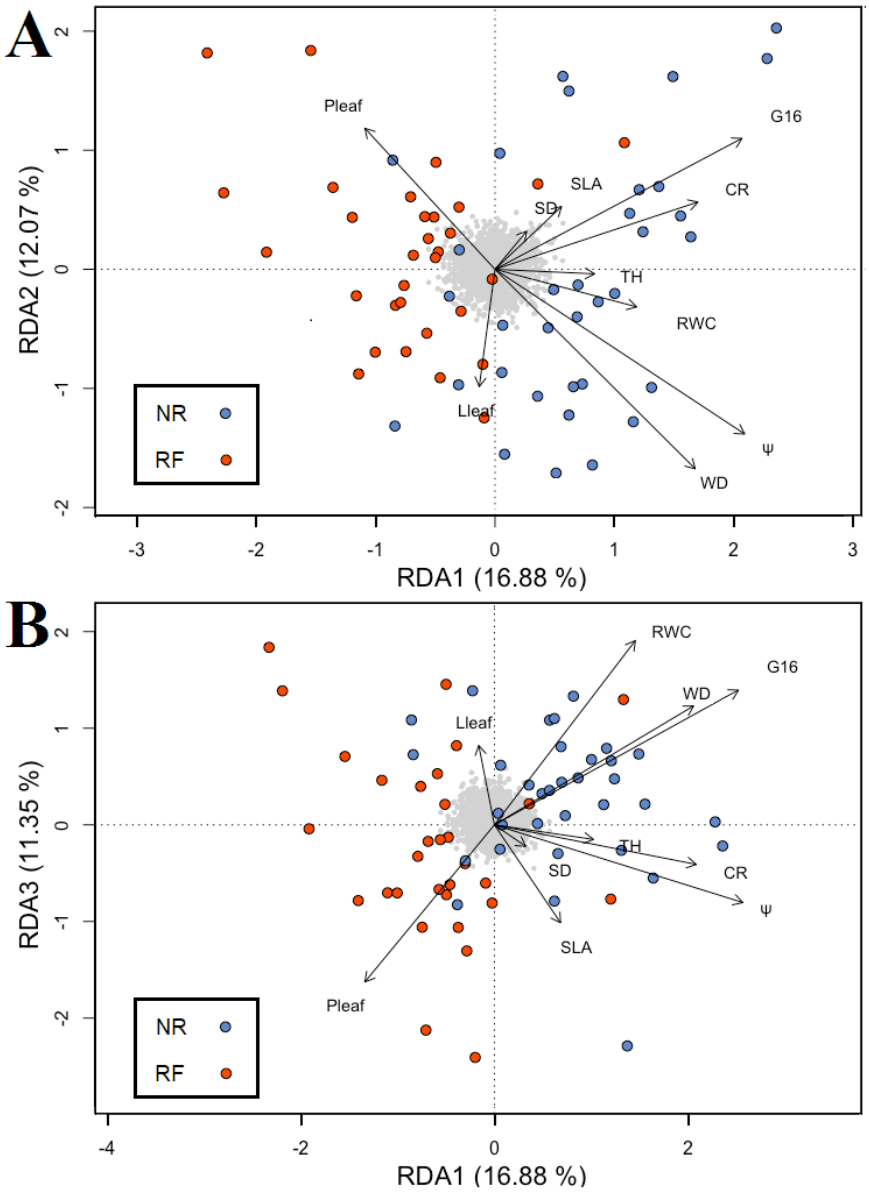

Fig. S6 Redundancy analysis (RDA) performed on genomic data from sacred fir (*Abies religiosa*) individuals sampled in a trial site in central Mexico. RDA performed with 1,749 filtered SNPs in $Model_{GxP3}$. Black arrows represent phenotypic traits: Total height (TH), stem diameter (SD), crown radius (CR), growth in 2016 year (G_{16}), WD (wood density), water potential (Ψ), relative water content (RWC), specific leaf area (SLA) and P concentration (P_{Leaf}). Dot color represents individual origin: blue, natural regeneration; red, introduced plants from reforestation with foreign germplasm. **A**. Components RDA1 and RDA2; **B.** components RDA1 and RDA3. *P*-value = 0.12.

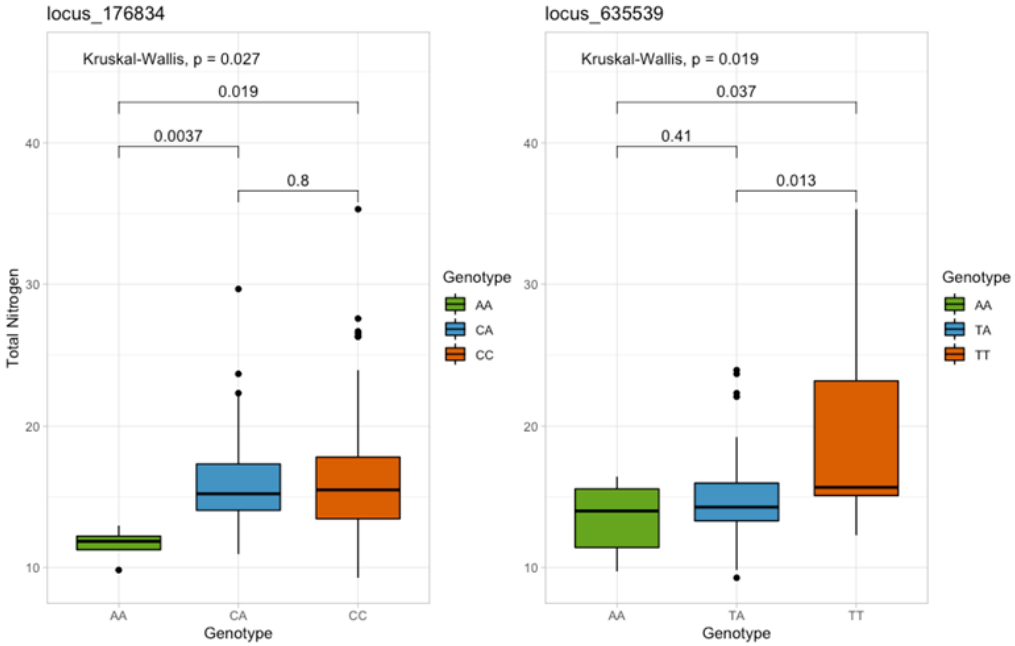

Fig. S7 Significant genotype – environment association for two annotated contigs containing candidate SNPs (*Locus*_176834 (left) and *Locus*_635539 (right)) in a sacred fir (*Abies religiosa*) trial site in central Mexico. Statistical significance according to LFMM results (Krustal-Wallis one-way; *P*-value <0.05).

CAPÍTULO III

¿Es posible utilizar modelos de selección genómica (SG) en poblaciones naturales de plantas con largos tiempos generacionales y aplicarlas en manejo y conservación?

¿Es posible utilizar modelos de selección genómica (SG) en poblaciones naturales de plantas con largos tiempos generacionales y aplicarlas en manejo y conservación?

Arenas S, Jaramillo-Correa JP

Resumen

Los estudios de mejoramiento genético se realizan solo desde hace pocas décadas en árboles, revelándose como una tarea lenta, complicada y costosa. Sin embargo, los avances en genómica y biología molecular han revolucionado las técnicas de genética cuantitativa tradicional aplicadas a datos familiares para cuantificar la heredabilidad de rasgos fenotipicos. Con ellas, es ahora posible estimar la proporción de la varianza fenotípica explicada por marcadores moleculares sin depender de un pedigrí estructurado. La selección de genotipos basada en estas técnicas busca sobre todo reducir los ciclos reproductivos en plantaciones comerciales, y queda la duda sobre su efectividad en poblaciones silvestres. La selección genómica (SG) estima los efectos de *loci* potencialmente ligados a rasgos cuantitativos (QTL) evaluados para cientos de individuos con cientos a miles de variantes alélicas distribuidas a lo largo de su genoma, permitiendo evaluar su valor reproductivo genómico (*GEBV*). La SG permite predecir rasgos complejos, como el crecimiento, el rendimiento y varias propiedades de la madera. Recientemente también se está validando su uso para predecir adaptabilidad de los individuos al ambiente (por ej. sequía) y a las plagas. Aplicado a poblaciones naturales, este enfoque podría orientar programas de manejo y conservación al predecir fenotipos pre-adaptados a un ambiente futuro. Revisamos el estado del arte actual de la SG en el mejoramiento de árboles y discutimos cómo se podrían aplicar estos modelos predictivos en poblaciones naturales. Destacamos particularmente los desafíos para su implementación, como el manejo de la incertidumbre sobre el parentesco de los individuos y la manera de integrar las interacciones genotipo-ambiente (G × E). Finalmente, discutimos las perspectivas para aplicar este método en programas de migración asistida y adaptación al ambiente, junto a posibles perspectivas tecnológicas como la utilización de métodos de fenotipado de alto rendimiento y de *Machine Learning*, o la utilización de variantes epigenéticas para mejorar los modelos predictivos. Esta revisión es entonces un llamado a futuros estudios empíricos para ampliar estas técnicas en el manejo y conservación de especies silvestres en su ambiente natural.

3.1 Introducción

El cambio climático (CC) es una amenaza que pone en riesgo la salud y la adaptación de todas las especies, particularmente las de larga vida como los árboles, afectando su crecimiento y supervivencia[1–3]. El manejo de la diversidad genética de los árboles es un paso fundamental para garantizar y preservar los servicios agroecosistémicos[4,5], utilizando metodologías amigables con el ambiente, como la transferencia de germoplasma[6]. También, se ha planteado utilizar herramientas genómicas para apoyar el mejoramiento y la adaptación al cambio climático, a través de la selección de materiales élite y la identificación de genes potencialmente adaptativos[7]. Los árboles son un recurso económico fundamental para la obtención de materias primas (alimento, fibra, resinas, biocombustible y madera), y mejorar la ganancia genética por unidad de tiempo para caracteres asociados a la producción de estas materias primas es el objetivo final de su cría selectiva[6]. Esta se desarrolla por medio de una serie de pasos que buscan aumentar la frecuencia de alelos ventajosos en rasgos poligénicos como el crecimiento, la calidad de la madera, la producción de frutos y la resistencia a plagas[8–10]; lo que a su vez tiene el riesgo de dejar en segundo plano otros rasgos claves para mantener la dinámica de los ecosistemas, como la eficiencia en el uso del agua y la supervivencia[11,12].

Entre los desafíos para la cría selectiva en árboles se encuentran los largos tiempos generacionales de muchas especies, la baja correlación que existe entre los fenotipos de árboles juveniles y maduros, la variación en las respuestas a cambios temporales del ambiente y las presiones emergentes debido al CC[7]. Esto conlleva a un interés creciente por conocer los componentes genéticos detrás de sus respuestas a condiciones desfavorables[13,14]. El progreso en el conocimiento de la base genética de estas respuestas, acompañado de la necesidad de reducir las pérdidas económicas, han promovido la implementación de técnicas para seleccionar árboles con mayor potencial adaptativo o de mejora[7]. Por esto, el desarrollo metodológico de la selección genómica ha cambiado desde su origen aplicado en animales domésticos y se ha adecuado al estudio de las especies con ciclos de vida largos y con potencial para la conservación en ambientes naturales.

Aquí, nos centraremos en cubrir el estado actual de la optimización y aceleración del mejoramiento genético en árboles, explorando la posibilidad de extender estas técnicas al manejo y conservación de poblaciones naturales (PN). La idea central para esta exploración es apoyarse en datos genómicos (p. ej., SNPs; polimorfismos de un solo nucleótido) para orientar

los planes de reforestación, conservación, manejo de la diversidad y transferencia de germoplasma, con el objetivo de aumentar la supervivencia de las plántulas y reducir las pérdidas económicas[10]. Evaluamos si las estimaciones de la proporción de la variación fenotípica debido a factores genéticos (i.e. la heredabilidad) son un parámetro efectivo para hacer un puente exploratorio entre las investigaciones de genética cuantitativa tradicional y los métodos que incluyen marcadores moleculares. Para ello verificamos si existe correlación entre las estimaciones hechas con estas dos aproximaciones conceptuales. La razón detrás de esta hipótesis son los estudios que sugieren que los cálculos de *heredabilidad* efectuados a partir de caracteres fenotípicos y marcadores moleculares son a menudo similares, incluso entre poblaciones y ambientes diferentes[15]. Esto nos permitiría visualizar lo que podría suceder si se hace un uso generalizado de predicciones genómicas en poblaciones/individuos silvestres, como ya se ha hecho con plantaciones comerciales[7]. En segundo lugar, realizamos una breve revisión histórica sobre cómo y con qué fin se integran la genética cuantitativa tradicional y los estudios con marcadores para predecir fenotipos en árboles. Luego, describimos los desafíos para aplicar estos métodos en poblaciones naturales a la luz de algunos resultados prometedores[16–18]. Finalmente, intentamos echar un vistazo al futuro donde los avances progresivos de la selección o predicción genómica (SG y PG) proporcionarían una plataforma poderosa para realizar estudios adaptativos y predictivos ante el cambio climático.

3.2 El cambio climático y la importancia de los árboles

El cambio climático está provocando variaciones en el sistema climático de la Tierra debido a la acumulación de gases de efecto invernadero[19]. Se prevé que los regímenes climáticos para los próximos 100 años estén caracterizados por un aumento de la temperatura, la concentración de CO_2 en la atmósfera y la frecuencia de eventos extremos (p. ej., sequías prolongadas, huracanes)[14,20]. Todo esto debería impactar en el rendimiento y productividad de los ecosistemas, poniendo en riesgo la capacidad adaptativa de las especies, especialmente las de larga vida, como los árboles[1,21]. Los árboles juegan un rol importante en el complejo equilibrio biológico del planeta, ya que aportan servicios agro-ecosistémicos que facilitan la recarga de acuíferos y el secuestro de carbono. También reducen la degradación del suelo y ayudan a mitigar la inseguridad alimentaria en diversas poblaciones humanas[19,21]. De esta forma, y dada la lenta migración y adaptación genética de los bosques a largo plazo, se anticipa

que estos no puedan mantener el ritmo de un cambio ambiental tan acelerado[3,11]. Para intentar predecir cómo responderán los árboles a este cambio, se necesita conocer las restricciones que afectan su crecimiento y los mecanismos genéticos y funcionales claves involucrados en la respuesta a estas restricciones[22]. Estas brechas de conocimiento dificultan el pronóstico de las respuestas de los bosques a la amenaza del CC, y llaman a implementar enfoques integrados de gestión del paisaje[23] y prácticas de manejo y conservación, como la migración asistida de individuos pre-adaptados que garanticen la estabilidad ecológica[24,25].

La adaptación local en especies arbóreas se ha documentado a través de una larga historia de experimentos de jardín común, compuestos por individuos de diferentes procedencias; en ellos se evalúan rasgos cuantitativos que reflejan su *fitness*, como el crecimiento y la supervivencia[26,27]. La rapidez del CC exige, sin embargo evaluaciones de este tipo en poblaciones naturales para investigar su sensibilidad a las presiones ambientales (eg.[28–30]). Dado que los jardines comunes son costosos de establecer y mantener, y debido al largo período que toma pasar de su establecimiento a las recomendaciones para el manejo, el desarrollo de análisis moleculares se ve como una aproximación para reducir los ciclos generacionales y hacer predicciones fenotípicas más rápidas para la cría selectiva[13,31,32]. La utilización de herramientas genómicas ha ayudado a desentrañar la arquitectura de la adaptación local e identificar regiones génicas involucradas en las respuestas adaptativas, mismas que buscan ser integradas en el manejo y conservación de poblaciones naturales[1,28,33] Un ejemplo son las especies forestales que usualmente se distribuyen a lo largo de clinas climáticas, para las que se han observado modificaciones adaptativas a lo largo de estos gradientes, como el rebrote en respuesta a la altitud[34] o la presencia de cutículas foliares gruesas, que les permiten retener humedad y las protegen de la luz ultravioleta[35].

3.3 La heredabilidad: un parámetro ideal en las perspectivas de la selección genómica (SG) en poblaciones naturales (PN)

Para descifrar la arquitectura genética de un rasgo complejo, es decir, cuántos y cuáles *loci* influyen en este, y cómo están distribuidos sus efectos en el fenotipo (*size effects*), es importante comprender y predecir la evolución de dicho rasgo[16]. En este sentido, la evolución se entiende como el cambio de las frecuencias alélicas de dichos *loci* por efecto de la selección natural[36,37]. Evaluar el efecto de las fuerzas selectivas es necesario para predecir las

consecuencias de los cambios ambientales en la adaptación local, la conservación de la diversidad y la idoneidad del germoplasma a ser transferido (ej.[25,38]). Antes de que la información de marcadores moleculares estuviera disponible, los estudios de genética cuantitativa se basaron principalmente en modelos lineales de efectos mixtos (MLM), aplicados a datos fenotípicos familiares[15,39]. Estos estudios suponían que los rasgos estaban afectados por un número infinito de genes o *loci* asociados a rasgos cuantitativos (QTL) con efectos pequeños; estos permitían modelar las (co)variaciones genéticas de los rasgos y su cambio evolutivo, para luego hacer predicciones de los valores genéticos de los individuos a partir de la información de su pedigrí[40,41]. Con esta suposición, se desarrollaron estudios que cuantificaban el potencial adaptativo a partir de estimaciones de heredabilidad del fenotipo[37,42,43]. La heredabilidad (h^2) es un parámetro clave e intrínseco de las poblaciones[44], que mide la importancia relativa de la variación genética aditiva (σ_g^2) con respecto a la ambiental (σ_e^2), para explicar la variación fenotípica (σ_p^2) de un caracter[15,45]. Sin embargo, h^2 difiere entre los rasgos y poblaciones evaluados, lo que se traduce a una menor precisión para aquellos que están más afectados por el ambiente que por factores genéticos[16,43].

Recientemente, con el desarrollo de las técnicas moleculares avanzadas (como la secuenciación de nueva generación), evaluar el potencial adaptativo en organismos no modelo se ha vuelto cada vez más factible. En estos estudios, se identifican cientos a miles de variantes alélicas relacionadas con un rasgo de interés o 'modelo instrumental'. El modelo instrumental (p. ej. modelo de aprendizaje virtual 'GBLUP'[46]) estima la proporción de la σ_p^2 explicada por una regresión lineal en un panel de (cientos a miles de) marcadores; a esta varianza se le conoce como heredabilidad genómica (h_g^2)[15,45]. Estas técnicas también permiten el uso de métodos sofisticados, como la regresión de todo el genoma (WGR)[47], que está basada en el modelo instrumental y tiene un poder estadístico superior al MLM. La WGR se está usando cada vez más para predecir rasgos cuantitativos o categóricos en diversas especies[48], brindando la oportunidad de hacer predicciones precisas de la h_g^2 sin depender de un pedigrí estructurado. Esto se hace a partir de un gradiente de parentesco de relaciones genómicas aditivas entre los individuos analizados (matriz *G*), lo que disminuye los intervalos generacionales[15,45,46]. Sin embargo, debido a las diferencias entre los métodos de genética cuantitativa tradicional y la WGR, es esencial verificar si existe un vínculo entre ambos, para así entender en qué medida el genotipo y el ambiente pueden explicar una gran proporción de la varianza fenotípica (σ_p^2)

entre y dentro de las poblaciones de interés. Sin embargo, el número de reportes en plantas longevas utilizando enfoques cuantitativos clásico supera con creces a los de marcadores, y ambas aproximaciones están sesgadas hacía ciertas familias vegetales como Pinaceae, Myrtaceae y Rosaceae (Tabla 1).

3.4 Revisión bibliográfica de valores de heredabilidad en árboles

Realizamos una extensa revisión de la literatura para localizar publicaciones con ecuaciones predictivas que estiman la heredabilidad en sentido estricto (i.e. varianza aditiva) en plantas con ciclos de vida largos. Buscamos artículos científicos en motores de búsqueda como *Science Direct*, *Google Academic*, *Redalyc* y *Scopus*. Adicionalmente, hicimos una búsqueda de "literatura gris", como tesis de grado o reportes internos no publicados; en ambos casos, usando palabras clave con operadores lógicos para seleccionar la literatura relevante: "*narrow-sense heritability*, *growth performance*, *linear mixed model*, *genomic selection*, *genomic prediction*, *additive genetic variance*, *heritability and SNP*, *variance components*, *forest tree*, *frutal tree* y *tropical tree"*. Una vez localizado un documento que contenía una estimación de h^2, extrajimos la información para incorporarla en una base de datos con diversos campos: especie, familia, rasgo, enfoque de estudio; para la h_g^2, registramos el tipo y número de marcadores utilizados. Para los estudios que presentaron más de un valor por rasgo y por especie, promediamos los datos disponibles, descartando los valores extremos (cuando los hubo).

Los datos recopilados se analizaron con los paquetes *devtools*, *ggpmisc* y *ggplot2* versión 3.2.2[49] dentro del programa R[50]. Se registraron valores pareados para las especies de las que se obtuvieron valores de h^2 y h_g^2. Con ellos se elaboró una regresión lineal y se calculó la proporción de la heredabilidad perdida, definida como la diferencia entre h_g^2 y h^2 y estimada con la fórmula $(h^2 - h_g^2) / h^2$, para los rasgos más documentados en la literatura con el objetivo de identificar si los resultados entre las dos aproximaciones difieren en la proporción de σ_p^2 retenida. Para las especies con más de un estudio reportando alguna de las heredabilidades, se seleccionó aleatoriamente uno de los valores, mientras que para aquellas en las que hubo más de dos reportes de ambas aproximaciones, se seleccionaron dos estudios de cada tipo, también de forma aleatoria; esto con el objetivo de aumentar el tamaño de muestra del análisis.

Inicialmente, las regresiones lineales revelaron una correlación entre h^2 y h_g^2 para explicar la σ_p^2 reportada (R^2 = 0.22, 0.31, 0.12 y 0.17 y *P-values* = 0.004, 0.001, 0.2 y 0.18; Fig. 1) para los caracteres más estudiados en árboles: altura, diámetro, densidad de la madera y rendimiento. Dentro de la base de datos revisada, se concluye que los dos rasgos más analizados en las mismas especies de plantas por las dos aproximaciones (genética cuantitativa tradicional y modelo instrumental) fueron la altura total y el diámetro del tallo.

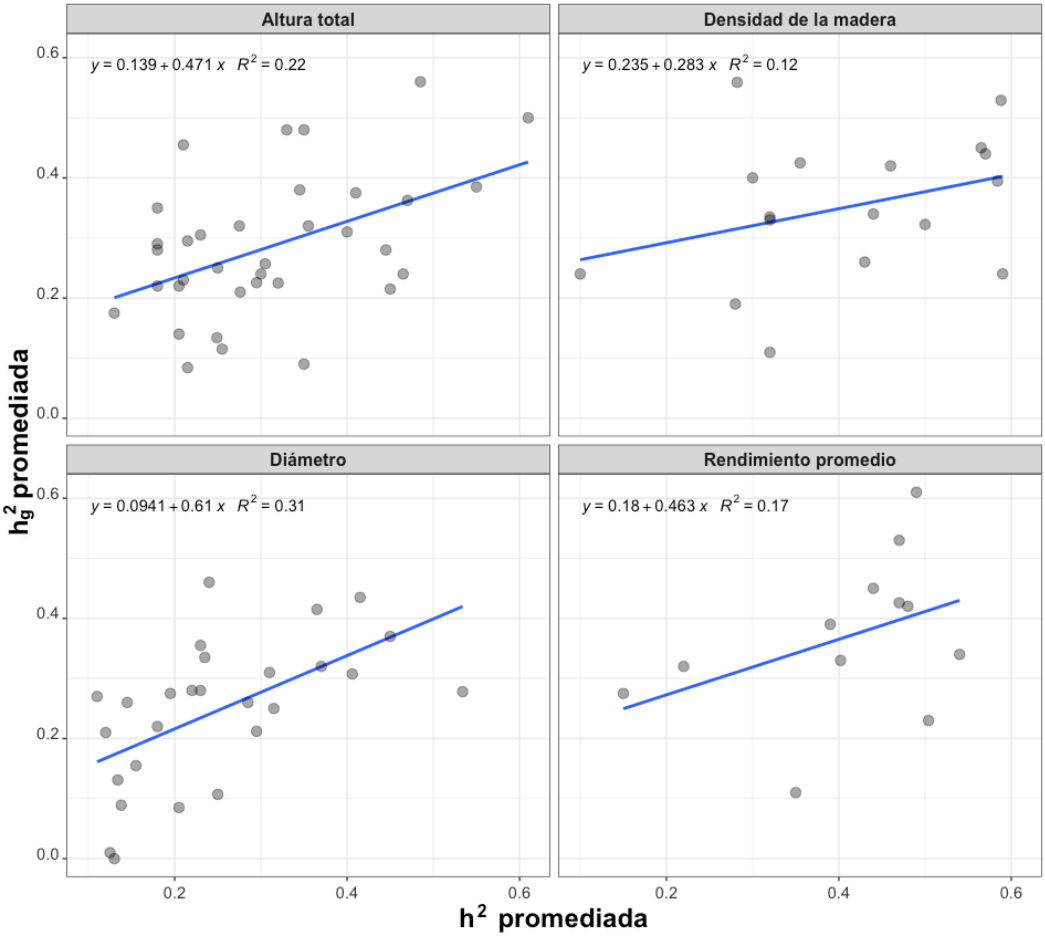

Fig. 1 Correlación entre los promedios de h^2 *vs.* h_g^2 para la altura, el diámetro, la densidad de la madera y el rendimiento promedio estimado en las mismas especies. Las líneas representan la ecuación que ajusta una regresión lineal y el valor de R^2 corresponde al coeficiente de correlación de Pearson.

Conceptualmente la inconsistencia debida a la 'heredabilidad perdida' se atribuye a factores como un número insuficiente de variantes (moleculares) causales dentro del modelo, una alta variación en el desequilibrio del ligamiento (DL) entre los marcadores evaluados y el/los locus(*loci*) causal (QTL)[51] o una elevada interacción genotipo-ambiente (G × E), entre

otros[52]. En la Fig. 2, presentamos el cálculo de 'heredabilidad perdida' para los estimados de las dos heredabilidades medias (h^2 y h_g^2) obtenidas para los rasgos más estudiados en árboles (altura total y diámetro). Esta fue positiva para muchos de los estudios revisados, lo que indica que la heredabilidad obtenida con un enfoque convencional fue mayor a la estimada con métodos basados marcadores moleculares (i.e. izquierda de la línea o $h_g^2 < h^2$). Hubo sin embargo algunas excepciones para las cuales el cálculo fue negativo (p. ej. *Picea mariana*), sugiriendo que el método de marcadores reportó valores mayores al tradicional (i.e. derecha de la línea o $h_g^2 > h^2$). No obstante estas diferencias, los cálculos de h^2 y h_g^2 fueron relativamente similares (aunque en poblaciones diferentes) y confirman que ambos enfoques son probablemente complementarios y contribuyen cada uno a su manera a desenmarañar la arquitectura genética del fenotipo y a entender la evolución del mismo[45].

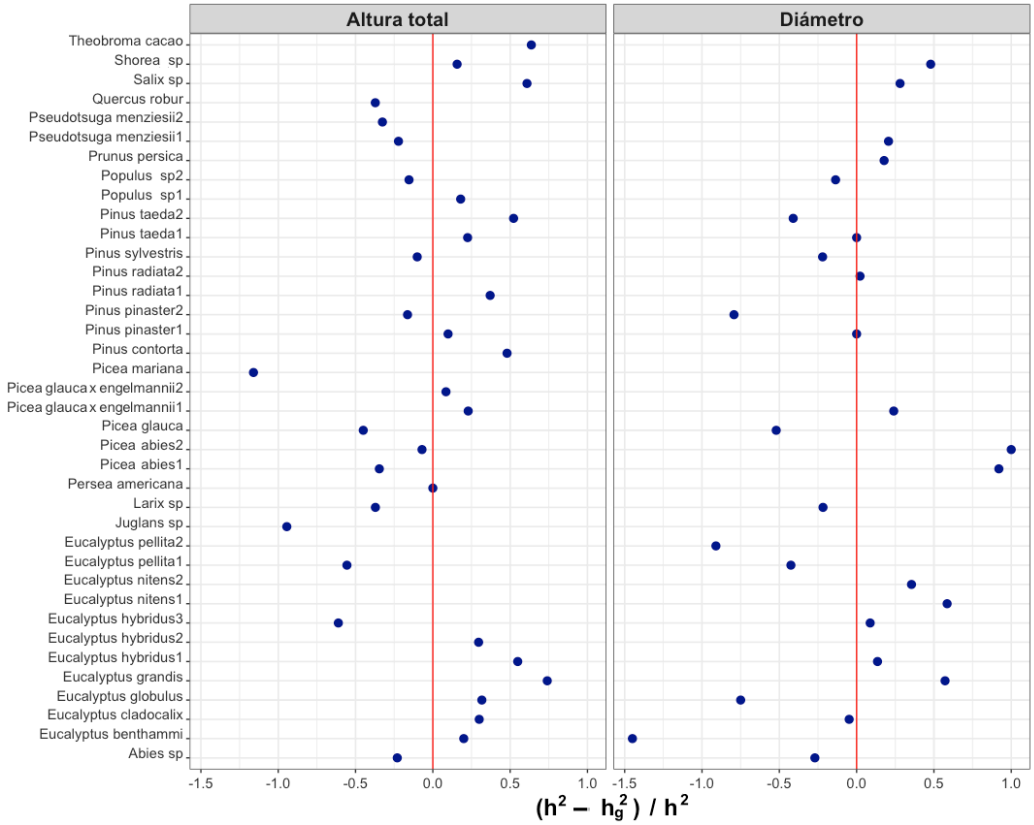

Fig 2. Estimaciones de heredabilidad perdida ((h^2 - h_g^2) / h^2) a partir de estudios para una misma especie de árbol. La línea roja representa la ausencia de la misma ($h^2 = h_g^2$).

3.5 El camino hacia la SG en poblaciones naturales (PN)

Los avances en las técnicas de biología molecular condujeron al desarrollo de los métodos de predicción utilizando aproximaciones WGR, los cuales ayudan a incrementar la eficiencia

de los programas de mejora en plantas[53]. Hasta hace poco, la selección asistida por marcadores (SAM) era la más utilizada en árboles[54,55]. Con ella se buscaba ubicar y estimar el efecto de QTLs individuales, la mayoría de ellos con grandes efectos fenotípicos[41,55] para rasgos simples o binarios y en diferentes poblaciones/ambientes[56]. Aunque, el SAM daba información valiosa sobre la respuesta adaptativa, esta se limitaba a especies con genomas grandes, tiempos generaciones largos, recursos genómicos limitados y rasgos extremadamente poligénicos o con expresión tardía, como la fenología de las yemas foliares, la calidad de la madera y la altura[57,58].

Dadas las deficiencias arriba descritas, se comenzaron a realizar mapeos de asociación en genomas "completos" (GWAS) en varias especies[59,60], para fortalecer las técnicas de mejoramiento genético. Estos hacían estimaciones de la asociación 'marcador-rasgo' entre individuos[59,61], proveyendo una mayor información sobre la arquitectura de los rasgos. Estos estudios resultaron en largas listas de genes candidatos potencialmente adaptativos e influenciados por el ambiente (ej. temperatura, sequía y patógenos)[23,62,63]. Un caso bien documentado es el de las proteínas de choque térmico (*Hsps*), que podrían estar involucradas en la resistencia al estrés por temperatura y sequía[64]. Sin embargo, al igual que con el SAM, los GWAS retuvieron pocas variantes de baja frecuencia y con efectos fenotípicos modestos. También generaron una gran cantidad de falsos positivos, es decir genes neutrales que fueron identificados como candidatos[65].

La falta de poder predictivo de los GWAS para varios rasgos provocó un cambio de paradigma y enfoque técnico. Se sugirió entonces que la selección y predicción genómica (SG y PG) podría ser una buena herramienta para asociar las técnicas de genética cuantitativa clásica y la estimación de asociaciones discretas, ya que permite explorar miles de marcadores distribuidos a lo largo del genoma y estimar valores genómicos de reproducción (*GEBV*) de una manera más precisa[47]. Esto a su vez aceleraría las mejoras fenotípicas de rasgos complejos[7,8,31,66]. La SG fue diseñada para explorar modelos poligénicos en rasgos complejos, incluso para aquellos con baja h_g^2, como los caracteres de historia de vida, la fisiología y la interacción genotipo - ambiente (G × E)[17,67,68]. Las investigaciones en SG solo fueron posibles después del desarrollo de tecnologías de secuenciación masiva, que permitieron incluir un gran número de variantes (p. ej. SNP) que capturen proporciones sustanciales de h_g^2 [9,69], en lugar de unos pocos marcadores potencialmente causales[70]. Por ejemplo, diversos estudios utilizaron

material genético georreferenciado (diferentes procedencias) plantado en entornos comunes, para estimar la efectividad de la predicción del fenotipo, especialmente del crecimiento, la calidad de madera y la resistencia a enfermedades (ej. [17,71,72]). Esto resultaría en un mayor potencial para pronosticar el rendimiento de los individuos y preseleccionar los más adecuados para un plan de manejo específico, lo que aumentaría la ganancia genética y disminuiría los ciclos de selección[7,73]. Esta práctica se ha concentrado, hasta el momento, en plantaciones de algunas especies "modelo" de climas templados (p. ej. *Pinus taeda* y *Picea abies*) y para caracteres comerciales. Falta entonces explorar su potencial en especies menos estudias, particularmente tropicales, en entornos naturales y para caracteres ecológicamente importantes, como la tasa fotosintética, el potencial hídrico, la adaptación al suelo, la ganancia de carbono, la supervivencia, etc[74,75].

El hecho de que la heredabilidad genómica refleje las relaciones entre individuos[15,45], que haya estimaciones similares de h_g^2 en el campo y el laboratorio para una misma especie[15], y que haya ajuste entre los valores de h_g^2 y h^2 (ver sección anterior), nos llevan a pensar que la SG podría ayudar a predecir fenotipos individuales en poblaciones naturales, al menos en algunos casos muy específicos (p. ej. poblaciones históricamente pequeñas y aisladas, especies clonales etc.). Esto se argumenta de la siguiente manera: *i*) la poblaciones naturales podrían 'imitar' a las poblaciones parentales de un programa de mejoramiento, en el caso en que estas estén sobre todo compuestas principalmente por parientes que comparten alelos[67] (como en poblaciones aisladas de alta montaña). Esto se basa en el hecho de que la SG explota al máximo las relaciones de parentesco entre individuos, omitiendo el pedigrí, incluso en especies de polinización abierta[76,77]. *ii*) Las poblaciones naturales históricamente pequeñas y aisladas pueden tener un alto DL entre SNPs y QTLs, dado que sus tamaños efectivos históricos se presumen bajos, lo que resulta en una alta consanguinidad (parientes cercanos) y en niveles reducidos de recombinación[78]. Esto a su vez facilitaría la implementación de la SG en estas poblaciones. *iii*) Existen numerosos reportes con exactitudes de predicción aceptables usando una baja densidad de marcadores (~10K SNPs; Tabla 2), lo que haría posible integrar estos métodos en especies poco estudiadas y con información genética escasa, como los árboles tropicales o aquellos amenazados de extinción. *iv*) Estos métodos además pueden integrar de manera relativamente sencilla la interacción G × E en entornos múltiples[17,68], lo que facilitaría la aplicación de la SG a lo largo del rango de distribución de la especie de interés. Finalmente,

v) existen resultados de GWAS validados en ambientes naturales para muchas especies[60,79]; lo que indica que también se pueden hacer predicciones en estas condiciones y que los marcadores candidatos pueden ser incluidos en los modelos predictivos. De hecho, un estudio piloto en *Abies religiosa* argumenta la posibilidad de hacer predicciones en rasgos morfológicos y fisiológicos en poblaciones de ambientes naturales a partir de un amplio espectro de relaciones de parentesco genético[67].

Dado los resultados prometedores de este estudio piloto, describimos en la Fig. 3, las dos fases propuestas para implementar la SG en individuos/poblaciones naturales de árboles a largo plazo. En la primera fase se elaborarían y pondrían a prueba los modelos de predicción genómica (PG). Esta comienza con la determinación de un conjunto inicial de árboles de 'entrenamiento' (TRN), que son fenotipados y genotipados con un gran número de marcadores. A partir de estos datos, se desarrollan varios modelos predictivos que son posteriormente validados con un segundo conjunto de árboles, o de prueba (TST), que idealmente deberían estar (o se espera que estén) emparentados con los del TRN. Los modelos validados se utilizarán en la segunda etapa, que es donde la SG se pone realmente en práctica. En poblaciones naturales, el TRN estaría compuesto idealmente por cientos o miles de individuos (sin cruces controlados) con edades relativamente similares y muestreados en el mismo año y condiciones; estos son los que van a 'imitar' la población parental en un programa de mejoramiento. Lo esperado es que mientras más grande sea el número de estos individuos, mejor será la estimación de los efectos de cada marcador y más robusto será el modelo predictivo. En la segunda fase, los modelos de predicción se aplican en los individuos 'candidatos a selección'; por ejemplo, aquellos con los que se pretende reforestar o conservar un área determinada. A ellos se le determinarán los genotipos y con ellos se predecirán sus *GEBV*[7,66]. Finalmente, los individuos mejor clasificados con respecto su *GEBV* serían los individuos a integrar en los planes de reforestación, manejo o conservación; estos serían plantados y entrecruzados para producir la siguiente generación e iniciar un nuevo ciclo de SG.

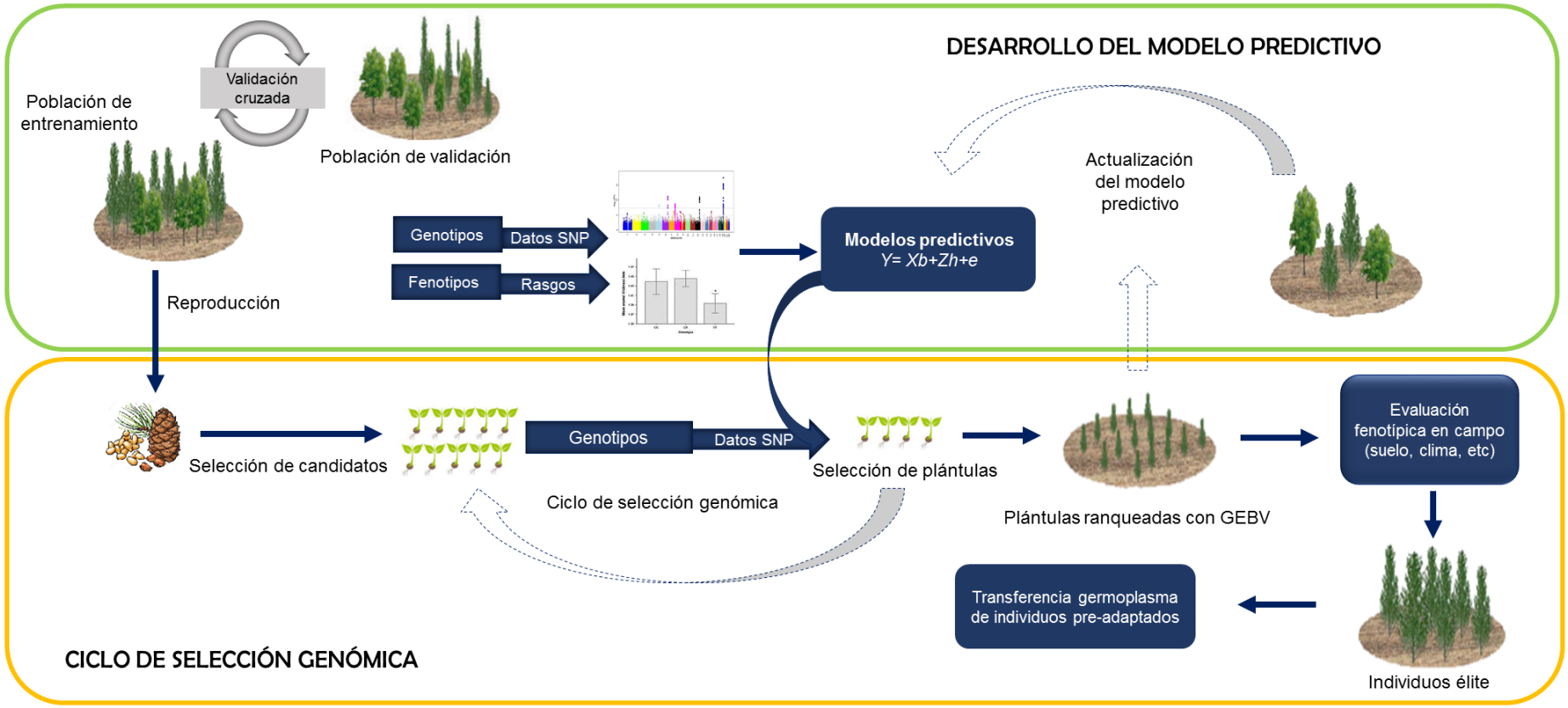

Figura 3. Esquema para integrar la selección genómica (SG) en poblaciones silvestres de árboles. Se inicia con el desarrollo de un modelo predictivo para los rasgos de interés (panel superior), éstos se utilizan para los ciclos de selección (panel inferior) y se retroalimentan con los datos generados. La SG utilizaría marcadores distribuidos a lo largo del genoma para estimar conjuntamente sus efectos en el fenotipo a partir de una "población de entrenamiento" (TRN) [Modificado de [10]]. Ver texto para una descripción detallada

3.6 SG: avances y grandes desafíos en poblaciones naturales de árboles

En la tabla 2 se muestra una lista de estudios empíricos en especies de árboles forestales y frutales que han calculado fenotipos a partir de modelos predictivos construidos con marcadores moleculares, la gran mayoría utilizando datos de genotipado de alto rendimiento. En ella se aprecia que hay más estudios para árboles forestales que frutales, y que la exactitud predictiva (*r*) para la mayoría de los rasgos está entre aceptable (estimaciones entre 0.2 y 0.4) y buena (estimaciones mayores a 0.4). Interesantemente, algunos de ellos son similares o mayores a los obtenidos en estudios genético-cuantitativos tradicionales, basados en pedigrí[80,81]. En las especies frutales se han evaluado principalmente rasgos relacionados con el rendimiento y la producción del fruto, y en las forestales los relacionados con el crecimiento y las propiedades de la madera. Aunque, se observa que muchos estudios tienen como objetivo el mejoramiento genético de una única población (*whole-sample*) con una estructura familiar conocida (principalmente medios hermanos[66]), recientemente se está empezando a trabajar sobre poblaciones con relaciones genéticas bajas o nulas, en individuos derivados de padres silvestres estudiados en ensayos de procedencia (p. ej.[18,82]) y en estudios en varias localidades, lo que permite medir los efectos de la interacción G × E. Por ejemplo, en la primera generación de árboles plus de *Cryptomeria japonica*, caracterizadas por relaciones genéticas muy bajas, se observó una exactitud de predicción mayor en los modelos *whole-sample* que para los de poblaciones individuales. Para los primeros se obtuvieron valores entre 0.193 para la densidad de madera y 0.634 para la fecundidad masculina[18]. Del mismo modo, en modelos diseñados para plantas de 19 procedencias, y sin ciclos de mejoramiento, se reportaron valores entre 0.05 para la lignina total y 0.30 para el volumen de la madera[82]. A partir de estos estudios y de algunas investigaciones que han realizado predicciones prometedoras para rasgos fisiológicos y de resistencia a plagas y enfermedades (p. ej.,[72,74]) incluso en especies clave con plagas naturales[83], faltaría confirmar qué tan bien se pueden transferir estos modelos a poblaciones naturales.

Uno de los grandes desafíos de la SG es tener en cuenta y explorar los factores que dificultarían su transferencia a poblaciones naturales, como el tamaño de muestra para el fenotipado y genotipado del conjunto TRN y la estructura genética subyacente, o la forma en que varían los rasgos entre individuos jóvenes y adultos[84,85], la presencia de interacciones G × E a diferentes escalas espaciales[86], el tipo de métodos para desarrollar y probar el modelo[87], la cantidad de marcadores que se incluyen en este[88] y la manera de evaluar su desempeño a largo plazo[89]. Algunos de estos desafíos ya se están investigando

en poblaciones de mejora genética con cruzas controladas[85,87,90,91]; explorarlos en poblaciones naturales sería entonces una de las prioridades para poder expandir satisfactoriamente la SG. A continuación, analizamos algunas ideas para explorar los tres primeros factores en poblaciones naturales a la luz de resultados experimentales en plantaciones.

3.6.1 Tamaño poblacional de la población de entrenamiento (TRN) y estructura genética

La aplicación de modelos de SG en árboles se basa en la estructura y diversidad genética de las poblaciones reproductoras; esta se usa para explicar las relaciones genéticas de parentesco entre el TRN y los individuos candidatos a selección. Así, un tamaño grande de entrenamiento garantizará una mayor confianza en las estimaciones, ya que tendría una representatividad más adecuada de las relaciones de parentesco (i.e. N_e [40]) y por ende de las relaciones genotipo-fenotipo. Esto produciría resultados con una probabilidad menor de estar sesgados. En general, los estudios de PG en árboles se han realizado entrenando los modelos con cientos a miles de individuos muestreados en diseños familiares, mostrando niveles de predicción aceptables para la mayoría de rasgos[90,92–94]. En poblaciones naturales y dependiendo de la especie, sería complicado tener conjuntos de entrenamiento grandes (por tanto una baja representatividad del N_e de la población), principalmente a causa de los desafíos metodológicos para fenotipificar tantos individuos. Tradicionalmente, los rasgos fenotípicos se determinan mediante mediciones manuales, a menudo destructivas, que requieren mucho tiempo y mano de obra. Muchas medidas son además propensas al sesgo durante su recopilación[66]; por ejemplo, para rasgos fisiológicos o aquellos relacionados la actividad metabólica, que pueden variar a lo largo del año o incluso durante un mismo día (p. ej., la conductancia estomática y el intercambio gaseoso[95]). Igualmente, algunos rasgos son complicados de medir para gran número de individuos en campo, como el potencial hídrico luego de un evento de estrés puntual, o la presencia de organismos simbiontes específicos (i.e. a nivel de especie) como micorrizas u hongos endófitos[37]. En algunos de estos casos, el fenotipo se ha logrado abordar utilizando métodos indirectos (*proxies*), por ejemplo el área foliar específica (*SLA*) y la tasa neta de asimilación (*NAR*) se han utilizado como medida indirecta de la ganancia de CO_2 en plantas[96,97].

El desarrollo de métodos de fenotipado de alto rendimiento (*HTP*), que son parte integral de la fenómica vegetal y se basan en imágenes de diversas regiones espectrales para su posterior procesamiento informático[98], han ayudado a solucionar, al menos

parcialmente, varios de los problemas arriba discutidos. Por ejemplo, han permitido deducir cambios fenotípicos después de algún evento ambiental. Estos métodos permiten un fenotipado preciso y no invasivo de numerosas plantas, y por largos períodos de tiempo, que se pueden efectuar para toda la planta (holísticos) o para órganos específicos[99]. A la par, también se pueden usar para medir rasgos de difícil fenotipificación, como la tolerancia a factores bióticos y abióticos. Por ejemplo, se ha simulado la infección artificial de plagas y calculado la eficiencia en el uso del agua (WUE)[74] usando espectroscopía de infrarrojo cercano (*NIRS*) de isótopos de carbono estable ($\delta^{13}C$) en híbridos de *Eucalyptus urophylla* × *E. grandis*. También se han implementado *proxies* para medir la resistencia a plagas, como la concentración de piceol y el pungenol, que confieren resistencia a la polilla *Choristoneura fumiferana* en abetos (*Picea*)[72].

A parte de la obtención de fenotipos, en poblaciones naturales es difícil conocer el parentesco entre los individuos, por lo que es necesario realizar análisis genético-poblacionales previos para estimarlo, por ejemplo a partir de matrices G[100]. Dado que los modelos de PG no funcionan muy bien entre individuos poco relacionados o entre poblaciones diferenciadas genéticamente, es indispensable tener de antemano una idea de este factor antes de iniciar la modelización. En la Fig. 4 presentamos un análisis de componentes principales (PCA) extraído de[101]. Se observa que la diferenciación genética es débil entre poblaciones, es decir que no se observa separación de estas, lo cual es importante para generar buenos modelos de predicción fenotípica. Dicho de otra manera, es más fácil predecir individuos genéticamente similares. Por ejemplo, se reportaron valores predictivos de entre 0.02 y 0.46 para un modelo de dos poblaciones conjuntas de *Eucalyptus nitens* con diferentes patrones de selección y una diferenciación genética parcial, cuando se realizaron modelos cruzados (i.e. entre poblaciones) estos valores decayeron a -0.07 y 0.08[102]. Por lo tanto, sería importante asegurarse de trabajar en poblaciones con bajos coeficientes de endogamia y reducida estructura genético-poblacional para tener modelos más precisos y realizar predicciones entre entornos.

Al igual que con el parentesco, también se debe determinar el desequilibrio de ligamiento (DL) promedio entre marcadores, y calcular la diversidad genética para tener una idea de la factibilidad de los modelos en la población de interés. El valor de DL va a garantizar que se esté trabajando con el tamaño de muestra óptimo (N_e), asegurando confiabilidad en los cálculos. En caso de querer trabajar con más de una población, también es recomendable contar con información *a priori* de la especie, a través de estudios de filogeografía y genética cuantitativa, ya que los modelos de PG no funcionan

adecuadamente en poblaciones con historias filogeográficas distintas. Un estudio de filogeografía garantizará que se estén utilizando poblaciones genéticamente similares y con historias demográficas equivalentes, que puedan reflejar la presencia de alelos compartidos entre poblaciones[67]. Generalmente, en el caso de encontrar que las poblaciones están muy diferenciadas entre sí, lo que se acostumbra hacer en cría selectiva es realizar modelos predictivos del total de individuos (p. ej.[90]), para luego hacer cruzas y transferencias recíprocas en caso de que se obtengan resultados importantes (p. ej. una alta h_g^2). Teniendo en cuenta los resultados que se muestran en la Fig. 1, contar con información *a priori* de la h^2 a partir de enfoques tradicionales servirá para darse una idea de cuáles rasgos vale la pena predecir y cuales no (además de que ya se tendría un experimento montado).

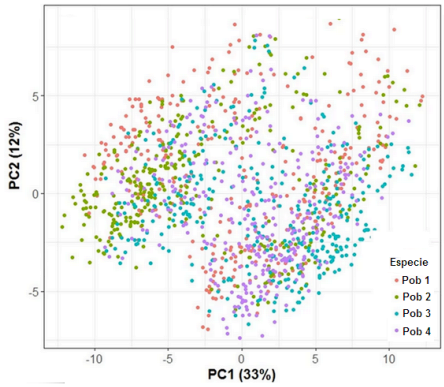

Fig 4. Análisis de componentes principales (PCA) que representa un caso ideal para el desarrollo de modelos de predicción genómica (La figura fue modificada del estudio de [104]), donde la diferenciación genética es débil entre cuatro poblaciones de una especie (ver paleta de colores).

3.6.2 Variación de los rasgos con la edad

A diferencia de las plantas anuales, como la mayoría de cultivos agrícolas, la edad en la que se realiza la fenotipificación es importante para hacer buenas predicciones genómicas en árboles. Dado que la correlación entre lo observado en árboles jóvenes y maduros puede cambiar significativamente según el rasgo, la especie y el entorno[7]. De acuerdo con algunos datos publicados, las predicciones son más efectivas cuando los individuos del conjunto TRN y TST tienen aproximadamente la misma edad. Por ejemplo, al evaluar la altura de híbridos de *Picea glauca* × *P. engelmannii* de 3, 6, 10, 15, 30 y 40 años de edad, se observó que la predicción de los modelos fue mucho menor para los individuos más jóvenes (entre los 10 y los 15 años), que para los árboles de 30 y 40 años (cuyas predicciones fueron equivalentes)[85]. En contraste, se han reportado predicciones similares para las propiedades de la madera para *Picea abies* de entre 6 y 19 años[84]. En *Populus deltoides* se reporta un caso similar, al observarse una correlación aceptable (0.39–

0.42) entre la altura de plantas en invernadero (13-15 semanas de edad) con las medidas en campo a los 3- 5 años[32]. Estas comparaciones indican que puede haber rasgos más 'estables', y que no cambian con la edad, y que son ellos los que se deben privilegiar para desarrollar modelos predictivos, probablemente incluyendo algún componente ambiental como co-variable.

En la naturaleza, lamentablemente se desconoce el nivel de correlación entre plantas jóvenes y maduras para la mayoría de rasgos; aunque algunas soluciones serían: estandarizar los rasgos de interés corrigiendo por la edad de las plantas (ver p. ej. [103]), construir modelos para diferentes rangos de altura o diámetro[67] o simplemente utilizar el crecimiento anual relativo de los individuos. Es importante resaltar que los datos obtenidos para especies perennes en diferentes edades se han recopilado principalmente en plantas cultivadas y en condiciones controladas; por lo tanto no pueden ser extrapolados totalmente a lo que se observaría en la naturaleza, dado que allí las presiones ambientales son cambiantes[66]. Cabe anotar también, que en regiones tropicales puede ser difícil determinar la edad de los árboles, ya que allí, al no haber estaciones, los anillos de crecimiento no están correlacionados con los ciclos anuales, como para los árboles templados. Para ello se podrían utilizar métodos más sofisticados, como los isótopos estables, la estimación de la descomposición de la madera, o simplemente trabajar con la regeneración en claros con una edad conocida[104,105].

2.6.3 Interacción genotipo × entorno

Los experimentos en diversas condiciones ambientales son esenciales para evaluar la influencia del entorno en los rasgos fenotípicos, y comprender las interacciones entre el genotipo y el ambiente (G × E). Las interacciones G × E han sido bastante estudiadas en árboles y se utilizan para determinar la estabilidad de los genotipos en diferentes condiciones (micro)ambientales[86,92]. Desde los primeros estudios experimentales[90,92,102], se determinaron diferencias para un mismo rasgo entre individuos de diferentes entornos. Esto vendría a implicar que las predicciones genómicas son más efectivas cuando las TRN y TST están en un mismo sitio que cuando se elaboran entre sitios, pues se controla mejor el efecto G×E en el potencial predictivo del modelo. Así, se espera que las estimaciones de un modelo disminuyan cuando se usen entornos cruzados, es decir cuando entrenamos un modelo en una población (TRN) y lo intentamos validar en individuos muestreados en otro sitio (TST). Sin embargo, incluir la interacción G×E en los modelos PG haría más realista

la predicción de los fenotipos e incluso podría permitir hacer predicciones entre ambientes o para ambientes no muestreados.

En los árboles, la interacción G × E se ha estudiado en sitios con ambientes contrastantes[68]. Por ejemplo, para individuos de *Pinus taeda* evaluados en cuatro sitios, se encontró que el modelo dentro de un mismo sitio proporcionó mayor precisión de predicción (0.64–0.74) que el elaborado entre sitios (0.18–0.66)[106]. Esto también se observó en *Picea glauca* (predicciones dentro un sitio 0.53–0.83; entre sitios 0.23–0.72)[77]. Sin embargo, hay que tener en cuenta que el efecto G × E también puede depender de la naturaleza del rasgo. Por ejemplo, esta interacción parece más acentuada sobre el crecimiento que sobre las propiedades de la madera[87].

En poblaciones naturales únicamente se podría utilizar un modelo elaborado para un sitio (TRN) para intentar predecir los *GEBV*s en otro lugar (TST) si los entornos cruzados son muy similares y estables en el tiempo, de manera que los rasgos sensibles al ambiente (p. ej. los bioquímicos y la fisiológicos) puedan ser predecibles con mejor exactitud[7,74]. Es obvio que desarrollar modelos predictivos en ambientes silvestres representa un gran desafío y que sería complicado controlar todas las variables (bióticas y abióticas) que determinan los procesos G × E para asegurar que el entorno de validación sea similar al de entrenamiento. Sin embargo, también se podría intentar predecir el desempeño funcional en múltiples sitios y crear con ellos el modelo predictivo[1,24,25]. De hacerse correctamente, los modelos podrían integrarse para hacer manejo y mitigación ecológica, e impulsar la recuperación de bosques intervenidos[17,75,107].

3.7 SG: ¿una herramienta para la conservación en un futuro cercano?

Dada la cantidad creciente de herramientas genómicas disponibles para árboles, de los genes candidatos retenidos a través de estudios GWAS, de los valores aceptables o elevados de h^2 reportados en estudios tradicionales, y de los resultados prometedores en los estudios iniciales de PG (Tabla 2), se prevé una adopción más generalizada de estas herramientas para cumplir los objetivos originales del SAM en el mejoramiento y conservación de plantas[7,108]. Sin embargo, además de los desafíos arriba discutidos, implementar la SG en programas de conservación/manejo adaptativo de poblaciones naturales va a requerir mejorar el valor predictivo de los modelos y exigirá incluir co-variables ambientales (p. ej. utilizando bases de datos como ClimateBC[109] y NASA POWER[110]), como ya se ha hecho en cultivos[111]. La integración de estas co-variables es fundamental para describir de forma más precisa el desempeño de las poblaciones en

diferentes ambientes naturales. Esta integración será un primer paso para la evaluación de las poblaciones naturales en múltiples entornos, incluso en aquellos que no fueron medidos, e identificar sitios adecuados para introducir germoplasma[55,112].

Un área que habrá que explorar mejor es la elaboración de modelos PG para especies distribuidas en ecosistemas heterogéneos y megadiversos, como los bosques tropicales. En estos ambientes las especies enfrentan presiones selectivas bióticas y abióticas extremadamente variadas[33,113] y su variación fenotípica parece estar muy asociada al ambiente[29,114]. Esto las hace, a su vez, vulnerables a los cambios súbitos, como las olas de sequía[1,2]; lo que sugeriría que la SG sería una herramienta poderosa para la reproducción selectiva y manejo adaptativo de estos árboles, a través de programas más dinámicos. Por ejemplo, para predecir respuestas conjuntas a los regímenes de lluvias[115], la irradiación[116], la composición del suelo[117] y las interacciones bióticas[118]. Esto llevaría a proponer medidas de gestión forestal sostenible, manejo de la biodiversidad y transferencia de germoplasma de individuos seleccionados[38,119].

El caso contrario es el de los ecosistemas urbanos, que han presentado múltiples eventos de fragmentación del paisaje, reducciones en la diversidad, introducción de especies exóticas, aumento de la contaminación y cambios en la temperatura del aire[120]. Una aproximación utilizado SG podría aumentar la ganancia en patrones y procesos ecológicos; por ejemplo, seleccionando individuos más tolerantes a las condiciones urbanas para hacer forestación[121,122], utilizando criterios como el uso del suelo, la estructura de la vegetación, el porcentaje de cobertura del dosel y el uso de especies nativas. El enfoque de manejo de árboles pre-adaptados a estas condiciones podría ayudar a recuperar suelos degradados y mantener la estructura de las comunidades ecológicas urbanas[123].

Dado que hay especies para las cuales ya se ha determinado su importancia en el ecosistema, y se tienen estudios genético-poblacionales, se pueden desarrollar modelos PG para seleccionar individuos pre-adaptados para forestar y reforestar ecosistemas intervenidos[74]. Por ejemplo, buscando germoplasma resistente al estrés hídrico o nutricional, para recuperar y preservar las propiedades físicas del suelo, evitar la desertificación o el azolvamiento de los mantos acuíferos. Esto modelos también ayudarían a identificar rasgos con potencial de mejora. Algunos ejemplos de estas especies son la jacaranda (*Jacaranda copaia*[124]), el arrayán (*Luma apiculata*[125]), el mango (*Manguifera indica*[126]) o la guayaba (*Psidium guajava*[127]) y diferentes especies tropicales y templadas de sauce (*Salix* spp.[128,129]) y álamo (*Populus* spp.[130]).

Cabe anotar que para varias de las especies mencionadas anteriormente ya existe una gran cantidad información genómica, incluidos algunos genomas de referencia. Estos permitirían identificar e integrar otros tipo de variación en los modelos PG, como la epigenética (ej., la metilación y el control regulador de ARNs no codificante) y el número de copias de genes (CNV)[131,132]. El papel de los cambios epigenéticos en la variación del fenotipo es cada vez más estudiado y se ha evaluado en un contexto de cría selectiva de árboles[133]. Estos cambios parecen ser importantes en la adaptación, ya que ocurren mucho más rápidamente que los genéticos[134]. Dos ejemplos de ello son un estudio que revela la asociación entre la metilación del ADN y las condiciones climáticas de poblaciones naturales de *Quercus lobata*[135] y otro mostró una correlación entre los niveles de metilación del ADN y la producción de biomasa en híbridos de *Populus* sometidos a diferentes condiciones de irrigación[136]. Por su parte, algunos estudios en los CNV indican que otros tipos de variación como las translocaciones y duplicaciones cromosómicas están involucradas en efectos fenotípicos que no pueden ser capturados totalmente por los SNPs, y por tanto se incrementa así la heredabilidad perdida[131]. Aunque la mayoría de ejemplos vienen de animales, como el efecto del gen KIT en el pelaje del ganado[137] y la correlación entre el color y las condiciones ambientales[138].

Por otro lado, se espera que las técnicas de *Machine Learning* (ML) ayuden a hacer mejores predicciones de PG para ambientes y rasgos múltiples y sobre todo, para desenredar la varianza genética aditiva de los rasgos y los dominados por G × E[139]. Los desarrollos recientes de ML permitirán construir predicciones más precisas mediante la implementación de variables ambientales fusionadas, la diversidad de microhábitats dentro de una misma población y la divergencia genética entre poblaciones[75]. Estas técnicas permitirán desarrollar todo lo anterior en un contexto de adaptación de los bosques al cambio climático[140], y aprovechar la variación fenotípica de rasgos adaptativos de los árboles.

Finalmente, dado que las investigaciones iniciales de SG en árboles comenzaron con estudios de simulación[54] y dichos estudios aún se realizan ocasionalmente[66], cabe esperar un uso cotidiano de estas para predecir la factibilidad de la PG en poblaciones naturales. Esto será particularmente importante, pues la implementación de modelos predictivos en poblaciones naturales es costosa y prolongada, por los tiempos generacionales largos que tienen la mayoría de los árboles. Esto permitiría desarrollar estrategias óptimas en cada especie, rasgo y variación ambiental específica.

Tabla 1. Número de estudios revisados de genética cuantitativa tradicional y basados en marcadores moleculares que reportan valores de heredabilidad para rasgos fenotípicos en árboles entre (1981) y (2020).

Tipo de estudio	Familia	Cantidad de estudios	Porcentaje
Métodos de genética cuantitativa tradicional	Pinaceae	74	38.7
	Myrtaceae	37	19.4
	Rosaceae	13	6.8
	Euphorbiaceae	7	3.7
	Fagaceae	11	5.6
	Salicaceae	7	3.7
	Verbenaceae	5	2.6
	Lauraceae	5	2.6
	Otras	32	16.9
	Total	191	
Métodos de regresión de todo el genoma con marcadores moleculares	Pinaceae	33	29.7
	Myrtaceae	25	22.5
	Rosaceae	11	9.9
	Euphorbiaceae	8	7.2
	Fagaceae	4	3.6
	Salicaceae	9	8.1
	Malvaceae	4	3.6
	Otras	17	15.4
	Total	111	

Tabla 2. Estudios de selección genómica (SG) para varios rasgos fenotípicos en diferentes especies con tiempos generacionales largos utilizando diferentes modelos estadísticos y plataformas de genotipado de marcadores moleculares de NGS.

Especies	*Tamaño poblacional*	*Plataforma NGS*	*Numero de marcadores*	*Rasgos fenotípicos*	*Modelo de SG*	*Habilidad de predicción*	*Referencia*
				Importancia alimenticia			
Malus x domestica	1120	SNP array	7.692	Seis (PF)	rrBLUP, BL	0.68 a 0.89	[141]
Malus							142
Pyrus pyrifolia Nakai	765	SNP array	1.536	18 (C, PF y R)	BRR, BL, BA, BB, BCπ, GBLUP y RF	0.2-0.8	[143]
Citrus sinensi, C. reticulata, C. limon, C. aurantifolia, C. paradisi y C. máxima	676	SNP array	1.841	Diez y siete (C y PF)	RR, GAUSS, RF, BL, EN, BRR, BL, BA, BB y BC	0.05-0.94	[144]

Citrus sinensis x C. reticulata	180	DArT array	5.287	Diez (C y PF)	rrBLUP	0.53-0.64	[145]
Malus	278	SNP array	8,294	Doce (PF)	BLUP	-0.5 - 0.81	[146]
Coffea canephora	3570	GBS	38,106	Tres (C, P y R)	rrBLUP, BBR, BA, BB, BR, BVS, BCπ, BL, RF y fixedMLR	0 - 0,52	[147]
Theobroma cacao	1,345	SNP array	9640	Cinco (R)	GBLUP	0.07-0.48	[148]
	148	SNP array	3,733	Once (P y R)	GBLUP	0.37-0.67	[149]
	358	SNP array	390	Cinco (P y CM)	rrBLUP	0.09-0.34	[150]
Manihot	899	GBS	78,212	Cinco (P y CM)	GBLUP, RKHS, BL, RF, BA y BB	-0.02-0.6	[151]
	429	GBS	63,016	Venti ocho (R)	GBLUP, RKHS, BA, BBBC, BA, BL y RF	0,27-0,45	[152]
Musa spp	307	GBS	10,807	Quince (P, PF y R)	BRR, BL, BA, BB, BC, RKHS	0.39-0.73	[153]
				Importancia industrial			
Elaeis guineensis Jacq	112	SSRs y SNPs	135 y 46,933	Seis (P y CM)	rrBLUP, BA, BB, BCπ, BL, BRR, SVM y RF	0.19-0.30	[154]
	1,218	SNPs	20.000	Seis (P y CM)	rrBLUP, BA, B C, BRR y BL	0.43-0.75	[155]
Hevea brasilensis	435	SNP array	30,546	C	ABLUP, GBLUP y RHKS	0.2-0.8	[156]
Hevea brasilensis	330	SSR	332	Uno (Rendimiento)	BLR, RKHS, RR-BLUP	0.33-0.6	[157]
Fraxinus excelsior	1250	TruSeq (genoma completo)	100-50,000	Uno (R)	RR-BLUP	0.46 – 0.63	[83]
Jatropha curcas	78	DArT array	1,248	Dos (P)	rrBLUP, GBLUP, BBR, BA, BB, BCπ, BL y RKHS	0.46-0.66	[158]
Castanea dentata	1230	GBS	71,507	Cinco(R)	HBLUP, Bayes C	0.60 – 0.95	[159]
Criptomeria japónica	476	SNP array	32,036	Cinco (C, M. Reproductive)	GBLUP, RF, Bayes B	-0.1 – 0.63	[18]
Criptomeria japónica	578	SNP	3,034	Qince (C y M)	GBLUP, RF	0.18-0.56	[160]
Eucalyptus benthamii	505	SNP array	7,563	Tres (C)	ABLUP, GBLUP, BRR, BA, BB, BCπ y BL	0-0.17	[93]
Eucalyptus pellita	732	SNP array	12,483			0-0.44	[93]
Eucalyptus polybractea	468	SNP array	10,000, - 502,000	Siete (C, terpenos y contenido aceite)	ABLUP, GBLUP, Bayes B	0.08-0.79	[161]
Eucalyptus cladochalix	480	SNP array	3,879	(Siete (C, M y F) FLORES	Bayes A, Bayes B, BAyes C, BRR	Calcular como yo	[162]
Eucalyptus nitens	691	SNP array	9,627	Seis (C y Prop MAd)	BLUP, GBLUP	0.18-0.78	[163]
Eucalyptus nitens	691	SNP array	12,236	Doce (C y M)	rrBLUP, GBLUP	-0.06-0.46	[102]
Eucalyptus globulus	726	SNP array	3,914	Siete (C, M y R)	rrBLUP, BL, BA y BB	0.38 a 0.60	[164]
E. globulus	582	SNP array	60000	Cinco (C)	BA, BB, BL, rrBLUP y PCR	0,43-0,54	[165]
E. robusta	415	DartT array	2,919	Tres (C y M)	ENet, GBLUP, RKHS	0.02-0.30	[82]
E. grandis	1,575	SNP array	15,040	Siete (C y M)	GBLUP	0.76-0.92	[166]
(E. grandis, E. urophylla, E. globulus y sus híbridos)	820 y 920	DArT array	3,564 y 3,129	Cuatro (C y M)	rrBLUP	0.55 a 0.88	[92]

E. urophylla x E. grandis	768	SNP array	24,806	Tres (C y M)	GBLUP	0.41-0.75	[71]
E. urophylla x E. grandis	949	SNP array	41,304	Ocho (C y M)	GBLUP, rrBLUP, BL y RKHS	0.12-0.47	[80]
E. urophylla x E. grandis	949	SNP array	41,304	Seis (C y M)	rrBLUP	0.19-0.5	[167]
E. urophylla x E. grandis	1130	SNP array	3,303	Cuatro (C y F)	BLUP	0.63-0.9	[74]
Populus trichocarpa	369	SNP array	~1,7 millones	Cinco (C y F)	rrBLUP	0.09-0.25	[168]
Populus deltoides	453	SNP array	68,885	Altura	GBLUP	0.25-0.47	[32]
Pinus contorta	1,569	SNP array	19,584	Seis (C y M)	GBLUP, BB, BC	0.34-0.8	[169]
Pinus taeda	790–840	SNP array	4,852	Dos (C)	rrBLUP	0.63–0.75	[106]
	951	SNP array	4,853	17 (C, R y M)	rrBLUP, BA, B Cπ, y BL	0.17 a 0.51	[170]
	149	SNP array	3,406	Cuatro (C y CM)	rrBLUP	0.30 a 0.83	[171]
	165	SNP array	3,461	C	rrBLUP y GBLUP	0.37–0.74	[172]
	951	SNP array	4,825	Ocho (C)	GBLUP	0.66–0.86	[173]
	758	GBS	2729-3531	Tres (C y M)	rrBLUP	0.09-0.23	[168]
Pinus pinaster	661	SNP array	2500	Tres (C)	GBLUP, BRR y BL	0.38–0.55	[174]
	817	SNP array	4332	Tres (C y M)	GBLUP y BL	0.24–0.94	[175]
Pinus radiata	523	GBS	58,636	Cinco (C y M)	BLUP y GBLUP	0.39-0.65	[163]
Pinus radiata	981	SNP array	67,108	Cuatro (C y F)	GBLUP, rrBLUP y BL	0.47–0.70	[176]
Pinus sylvestris	744	GBS	8719	Siete (C y M)			[177]
Picea glauca	1694	SNP array	6,385	Doce (C y M)	rrBLUP, BRR y BL	0.33 a 0.44	[77]
	1748	SNP array	6,932	Cuatro (C y M)	rrBLUP y BL	0.52 a 0.79	[100]
P. glauca x P. engelmannii	1126	GBS	8,868–62,198	Siete (C y M)	rrBLUP y GRR	0.47–0.77	[81]
P. glauca x P. engelmannii	769	GBS	34,570–50,803	C	rrBLUP, GRR y BCπ	0.31–0.47	[85]
Picea abies	1370	exome capture	77,116	Cuatro (C y M)	GBLUP, RKHS,BRR y BL	0.39-0.81	[178]
	714	SNP array	3,914	Siete (C, M y R)	GBLUP, ABLUP,BRR,B C y TGBLUP	0.1-0.55	[17]
Picea glauca	578 - 1,310	SNP array	5,308	Siete (C y R)	GBLUP y ABLUP, BCπ	0.38-0.67	[72]
Picea abies	484	SNP array	130,269	Seis (C y M)	GBLUP, ABLUP	0.35-0.46	[84]
Picea abies Confirmar especie	856	SNP array	4,092	Cinco (C y M)	Metiendo ablup es menos GBLUP	0.61-0.74	[179]
Picea mariana	734	SNP array	4,993	Cuatro (C y M)	rrBLUP	0.74-0.86	[90]
Picea sitchensis							[180]
Pseudotsuga menziesii	1372	Captura de exoma	69,551	Seis (C y M)	rrBLUP y GRR	0.78-0.92	[181]
	1372	Captura de exoma	69,551	C	rrBLUP, GRR y BB	-0.15-0.92	[112]
	11,759	Captura de exoma	66,969	C	ssGBLUP, PBLUP	0.44-0.63	[89]
Shorte platyclados	356	DDRadSeq	5,900	Siete (C y M)	BL, BRR, BA, BB y BC	-0.08-0.26	[182]
Abies religiosa	201	GBS	2286	Diez (C, M y F)	BRR y RKHS	0.11 – 0.54	[67]

NOTA: rasgos de crecimiento C, características de calidad de madera M, características químicas de madera CQ, resistencia a enfermedades R; GBLUP, BLUP genómico; rrBLUP, regresión aleatoria BLUP; BL, Bayesian LASSO; BC, BayesC; BA, BayesA; GRR, regresión ridge generalizada; BRR, regresión ridge bayesiana.

3.9 Referencias

1. Isabel, N., Holliday, J. A. & Aitken, S. N. Forest genomics: Advancing climate adaptation, forest health, productivity, and conservation. *Evolutionary Applications* **13**, 3–10 (2020).
2. Fischer, E. M. & Knutti, R. Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. *Nature Climate Change* **5**, 560–564 (2015).
3. Hansen, J., Sato, M. & Ruedy, R. Perception of climate change. *Proceedings of the National Academy of Sciences of the United States of America* **109**, (2012).
4. Navarro-Garza, H., Santiago-Santiago, A., Musálem-Santiago, M. Á., Vibrans-Lindemann, H. & Pérez-Olvera, M. A. La Diversidad De Especies Útiles Y Sistemas Agroforestales. *Revistas Chapingo Seria Ciencias Forestales y del Ambiente* **XVIII**, 71–86 (2012).
5. Ellison, D. *et al.* Trees, forests and water: Cool insights for a hot world. *Global Environmental Change* **43**, 51–61 (2017).
6. Koskela, J. *et al.* Utilization and transfer of forest genetic resources: A global review. *Forest Ecology and Management* **333**, 22–34 (2014).
7. Grattapaglia, D. *et al.* Quantitative genetics and genomics converge to accelerate forest tree breeding. *Frontiers in Plant Science* **871**, 1–10 (2018).
8. Grattapaglia, D. & Resende, M. D. V. Genomic selection in forest tree breeding. *Tree Genetics and Genomes* **7**, 241–255 (2011).
9. Bhat, J. A. *et al.* Genomic selection in the era of next generation sequencing for complex traits in plant breeding. *Frontiers in Genetics* **7**, 1–11 (2016).
10. Varshney, R. K., Roorkiwal, M. & Sorrells, M. E. *Genomic selection for crop improvement: An introduction*. *Genomic Selection for Crop Improvement: New Molecular Breeding Strategies for Crop Improvement* (2017). doi:10.1007/978-3-319-63170-7_1
11. Diamond, S. E. *et al.* Heat tolerance predicts the importance of species interaction effects as the climate changes. *Integrative and Comparative Biology* **57**, 112–120 (2017).
12. Barghi, N., Hermisson, J. & Schlötterer, C. Polygenic adaptation: a unifying framework to understand positive selection. *Nature Reviews Genetics* **21**, 769–781 (2020).
13. Nystedt, B. *et al.* The Norway spruce genome sequence and conifer genome evolution. *Nature* **497**, 579–584 (2013).
14. Nadeau, C. P., Urban, M. C. & Bridle, J. R. Climates Past, Present, and Yet-to-Come Shape Climate Change Vulnerabilities. *Trends in Ecology and Evolution* **32**, 786–800 (2017).
15. Stanton-geddes, J., Yoder, J. B., Briskine, R., Young, N. D. & Tiffin, P. Estimating heritability using genomic data. 1151–1158 (2013). doi:10.1111/2041-210X.12129
16. Gienapp, P., Calus, M. P. L., Laine, V. N. & Visser, M. E. Genomic selection on breeding time in a wild bird population. *Evolution Letters* **3**, 142–151 (2019).
17. Lenz, P. R. N. *et al.* Multi-trait genomic selection for weevil resistance, growth, and wood quality in Norway spruce. *Evolutionary Applications* **13**, 76–94 (2020).
18. Hiraoka, Y. *et al.* Potential of genome-wide studies in unrelated plus trees of a coniferous species, cryptomeria japonica (japanese cedar). *Frontiers in Plant Science* **9**, 1–15 (2018).
19. Schneider, S. *et al.* Overview of Impacts, Adaptation and Vulnerability to Climate Change. *Impacts, adapation and vulnerability* 75–103 (2011). doi:10.1016/j.gloenvcha.2005.11.007
20. Breshears, D. D. *et al.* The critical amplifying role of increasing atmospheric moisture demand on tree mortality and associated regional die-off. *Frontiers in Plant Science* **4**, 2–5 (2013).
21. Lobell, D. B. & Gourdji, S. M. The influence of climate change on global crop productivity. *Plant Physiology* **160**, 1686–1697 (2012).
22. Aubin, I. *et al.* Tree vulnerability to climate change: Improving exposure-based assessments using traits as indicators of sensitivity: Improving. *Ecosphere* **9**, (2018).
23. Housset, J. M. *et al.* Tree rings provide a new class of phenotypes for genetic associations that foster insights into adaptation of conifers to climate change. *New Phytologist* **218**, 630–645 (2018).
24. Aitken, S. N. & Whitlock, M. C. Assisted Gene Flow to Facilitate Local Adaptation to Climate Change. *Annual Review of Ecology, Evolution, and Systematics* **44**, 367–388 (2013).
25. Aitken, S. N. & Bemmels, J. B. Time to get moving: Assisted gene flow of forest trees. *Evolutionary Applications* **9**, 271–290 (2016).
26. Lascoux, M., Glémin, S. & Savolainen, O. Local Adaptation in Plants. *eLS* 1–7 (2016). doi:10.1002/9780470015902.a0025270
27. Blanquart, F., Kaltz, O., Nuismer, S. L. & Gandon, S. A practical guide to measuring local adaptation. *Ecology Letters* **16**, 1195–1205 (2013).
28. Santos, A. S. & Gaiotto, F. A. Knowledge status and sampling strategies to maximize cost-benefit

ratio of studies in landscape genomics of wild plants. *Scientific Reports* **10**, 1–9 (2020).

29. Martins, K. *et al.* Landscape genomics provides evidence of climate-associated genetic variation in Mexican populations of Quercus rugosa. *Evolutionary Applications* **11**, 1842–1858 (2018).

30. Oddou-Muratorio, S., Davi, H. & Lefèvre, F. Integrating evolutionary, demographic and ecophysiological processes to predict the adaptive dynamics of forest tree populations under global change. *Tree Genetics and Genomes* **16**, 1–22 (2020).

31. Badenes, M. L., Fernández i Martí, A., Ríos, G. & Rubio-Cabetas, M. J. Application of genomic technologies to the breeding of trees. *Frontiers in Genetics* **7**, 1–13 (2016).

32. Alves, F. C., Balmant, K. M., Resende, M. F. R., Kirst, M. & de los Campos, G. Accelerating forest tree breeding by integrating genomic selection and greenhouse phenotyping. *Plant Genome* 1–13 (2020). doi:10.1002/tpg2.20048

33. Carvalho, C. S. *et al. Combining genotype, phenotype, and environmental data to delineate site-adjusted provenance strategies for ecological restoration*. *Molecular Ecology Resources* **55**, (2020).

34. Ingvarsson, P. K., García, M. V., Hall, D., Luquez, V. & Jansson, S. Clinal variation in phyB2, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European aspen (Populus tremula). *Genetics* **172**, 1845–1853 (2006).

35. Yakovlev, I. *et al.* An adaptive epigenetic memory in conifers with important implications for seed production. *Seed Science Research* **22**, 63–76 (2012).

36. Savolainen, O., Pyhäjärvi, T. & Knürr, T. Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proceedings of the National Academy of Sciences of the United States of America* **38**, 1752–1756 (2007).

37. Charmantier A., G. D. y K. L. Quantitative Genetics in the wild. *Current Genomics* (2014). doi:10.2174/138920212800543110

38. Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T. & Curtis-McLane, S. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* **1**, 95–111 (2008).

39. Clark M. M., J. Blangero, T. D. Dyer, E. M. S. & and J. S. SINSHEIMER. The Quantitative-MFG Test: A linear mixed effect model to detect maternal-offspring gene interactions. *Physiology & behavior* **176**, 139–148 (2016).

40. Lin, Z., Hayes, B. J. & Daetwyler, H. D. Genomic selection in crops, trees and forages: A review. *Crop and Pasture Science* **65**, 1177–1191 (2014).

41. Barton, N. H., Etheridge, A. M. & Véber, A. The infinitesimal model: Definition, derivation, and implications. *Theoretical Population Biology* **118**, 50–73 (2017).

42. Cornelius, J. P. Heritabilities and additive genetic coefficients of variation in forest trees. *Canadian Journal of Forest Research* **24**, 372–379 (1994).

43. Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era - Concepts and misconceptions. *Nature Reviews Genetics* **9**, 255–266 (2008).

44. Brommer, J. E. Whither Pst? The approximation of Qst by Pst in evolutionary and conservation biology. *Journal of Evolutionary Biology* **24**, 1160–1168 (2011).

45. de los Campos, G., Sorensen, D. & Gianola, D. Genomic Heritability: What Is It? *PLoS Genetics* **11**, 1–21 (2015).

46. VanRaden, P. M. Efficient methods to compute genomic predictions. *Journal of Dairy Science* **91**, 4414–4423 (2008).

47. Meuwissen, T. H. E., Hayes, B. J. & Goddard, M. E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. (2001).

48. Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* **42**, 565–569 (2010).

49. Wickham, H. ggplot2-Book09hWickham. doi:10.1007/978-0-387-98141-3

50. Harris, R. An Introduction to R. *Quantitative Geography: The Basics* **3**, 250–286 (2018).

51. Zaitlen, N. & Kraft, P. Heritability in the genome-wide association era. *Human Genetics* **131**, 1655–1664 (2012).

52. Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 1193–1198 (2012).

53. Van Eenennaam, A. L., Weigel, K. A., Young, A. E., Cleveland, M. A. & Dekkers, J. C. M. Applied Animal Genomics : Results from the Field. *Annu. Rev. Anim. Biosci.* **2**, 105–139 (2014).

54. Grattapaglia, D. & Resende, M. D. V. Genomic selection in forest tree breeding. *Tree Genetics and Genomes* **7**, 241–255 (2011).

55. Grattapaglia, D. *Genomics of plant genetic resources: Volume 1. Managing, sequencing and mining genetic resources. Chapter 26 Breeding Forest Trees by Genomic Selection: Current Progress and the Way Forward*. *Genomics of Plant Genetic Resources: Volume 1. Managing, Sequencing and Mining*

Genetic Resources (2014). doi:10.1007/978-94-007-7572-5

56. He, J. *et al.* Genotyping-by-sequencing (GBS), An ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Frontiers in Plant Science* **5**, (2014).

57. Neale, D. B. & Kremer, A. Forest tree genomics: Growing resources and applications. *Nature Reviews Genetics* **12**, 111–122 (2011).

58. Sork, V. L. *et al.* Putting the landscape into the genomics of trees: Approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics and Genomes* **9**, 901–911 (2013).

59. Müller, B. S. F. *et al.* Independent and Joint-GWAS for growth traits in Eucalyptus by assembling genome-wide data for 3373 individuals across four breeding populations. *New Phytologist* **221**, 818–833 (2019).

60. Mckown, A. D. *et al.* Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of Populus trichocarpa. *New Phytologist* **203**, 535–553 (2014).

61. Lotterhos, K. E., Yeaman, S., Degner, J., Aitken, S. & Hodgins, K. A. Modularity of genes involved in local adaptation to climate despite physical linkage. *Genome Biology* **19**, 1–24 (2018).

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

63. Duruz, S. *et al.* Rapid identification and interpretation of gene–environment associations using the new R.SamBada landscape genomics pipeline. *Molecular Ecology Resources* **19**, 1355–1365 (2019).

64. Prunier, J. *et al.* The genomic architecture and association genetics of adaptive characters using a candidate SNP approach in boreal black spruce. *BMC Genomics* **14**, (2013).

65. Finno, C. J., Aleman, M., Higgins, R. J., Madigan, J. E. & Bannasch, D. L. Risk of false positive genetic associations in complex traits with underlying population structure: A case study. *Veterinary Journal* **202**, 543–549 (2014).

66. Lebedev, V. G., Lebedeva, T. N., Chernodubov, A. I. & Shestibratov, K. A. Genomic selection for forest tree improvement: Methods, achievements and perspectives. *Forests* **11**, 1–36 (2020).

67. Arenas S, Cortés AJ, Mastretta-Yanes A., J.-C. J. Evaluating the accuracy of genomic prediction for the management and conservation of relictual natural tree populations. *Tree Genetics & Genomes* (2021). doi:10.1007/s10832-007-9069-7

68. Li, Y., Suontama, M., Burdon, R. D. & Dungey, H. S. Genotype by environment interactions in forest tree breeding: review of methodology and perspectives on research and application. *Tree Genetics and Genomes* **13**, (2017).

69. Desta, Z. A. & Ortiz, R. Genomic selection: Genome-wide prediction in plant improvement. *Trends in Plant Science* **19**, 592–601 (2014).

70. Rice, B. & Lipka, A. E. Evaluation of rr-blup genomic selection models that incorporate peak genome-wide association study signals in maize and sorghum. *Plant Genome* **12**, 1–14 (2019).

71. Resende, R. T. *et al.* Assessing the expected response to genomic selection of individuals and families in Eucalyptus breeding with an additive-dominant model. **119**, 245–255 (2017).

72. Beaulieu, J. *et al.* Genomic selection for resistance to spruce budworm in white spruce and relationships with growth and wood quality traits. *Evolutionary Applications* 2704–2722 (2020). doi:10.1111/eva.13076

73. Resende, M. D. V. *et al.* Genomic selection for growth and wood quality in Eucalyptus: Capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytologist* **194**, 116–128 (2012).

74. Bouvet, J. M. *et al.* Selecting for water use efficiency, wood chemical traits and biomass with genomic selection in a Eucalyptus breeding program. *Forest Ecology and Management* **465**, 118092 (2020).

75. Cortés, A. J., Restrepo-Montoya, M. & Bedoya-Canas, L. E. Modern Strategies to Assess and Breed Forest Tree Adaptation to Changing Climate. *Frontiers in Plant Science* **11**, (2020).

76. El-Dien, O. G. *et al.* 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**, 743–753 (2016).

77. Beaulieu, J., Doerksen, T., Clément, S., Mackay, J. & Bousquet, J. Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. *Heredity* **113**, 343–352 (2014).

78. Finlay, C. M. V., Bradley, C. R., Jane Preston, S. & Provan, J. Low genetic diversity and potential inbreeding in an isolated population of alder buckthorn (Frangula alnus) following a founder effect. *Scientific Reports* **7**, 1–8 (2017).

79. Guerrero, J., Andrello, M., Burgarella, C. & Manel, S. Soil environment is a key driver of adaptation in Medicago truncatula: new insights from landscape genomics. *New Phytologist* **219**, 378–390 (2018).

80. Tan, B. *et al.* Evaluating the accuracy of genomic prediction of growth and wood traits in two Eucalyptus species and their F1hybrids. *BMC Plant Biology* **17**, 1–15 (2017).

81. Gamal El-Dien, O. *et al.* Prediction accuracies for growth and wood attributes of interior spruce in space using genotyping-by-sequencing. *BMC Genomics* **16**, 1–16 (2015).

82. Rambolarimanana, T. *et al.* Performance of multi-trait genomic selection for Eucalyptus robusta breeding program. *Tree Genetics and Genomes* **14**, (2018).

83. Stocks, J. J. *et al. Genomic basis of European ash tree resistance to ash dieback fungus*. *bioRxiv* (2019). doi:10.1101/626234

84. Zhou, L. *et al.* Effect of number of annual rings and tree ages on genomic predictive ability for solid wood properties of Norway spruce. *BMC Genomics* **21**, 1–12 (2020).

85. Ratcliffe, B. *et al.* A comparison of genomic selection models across time in interior spruce (Picea engelmannii × glauca) using unordered SNP imputation methods. 547–555 (2015). doi:10.1038/hdy.2015.57

86. Cappa, E. P., Stoehr, M. U., Xie, C. Y. & Yanchuk, A. D. Identification and joint modeling of competition effects and environmental heterogeneity in three Douglas-fir (Pseudotsuga menziesii var. menziesii) trials. *Tree Genetics and Genomes* **12**, (2016).

87. Chen, Z. Q. *et al.* Accuracy of genomic selection for growth and wood quality traits in two control-pollinated progeny trials using exome capture as the genotyping platform in Norway spruce. *BMC Genomics* **19**, (2018).

88. Thistlethwaite, F. R. *et al.* Linkage disequilibrium vs. pedigree: Genomic selection prediction accuracy in conifer species. *PLoS ONE* **15**, 1–14 (2020).

89. Ratcliffe, B. *et al.* Inter- and Intra-Generation Genomic Predictions for Douglas-fir Growth in Unobserved Environments. *bioRxiv* 540765 (2019). doi:10.1101/540765

90. Lenz, P. R. N. *et al.* Factors affecting the accuracy of genomic selection for growth and wood quality traits in an advanced-breeding population of black spruce (Picea mariana). *BMC Genomics* **18**, 1–17 (2017).

91. El-Dien, O. G. *et al.* Multienvironment genomic variance decomposition analysis of open-pollinated Interior spruce (Picea glauca x engelmannii). *Molecular Breeding* **38**, (2018).

92. Resende, M. D. V. *et al.* Genomic selection for growth and wood quality in Eucalyptus: Capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytologist* **194**, 116–128 (2012).

93. Müller, B. S. F. *et al.* Genomic prediction in contrast to a genome-wide association study in explaining heritable variation of complex growth traits in breeding populations of Eucalyptus. *BMC Genomics* **18**, 1–17 (2017).

94. Minamikawa, M. F. *et al.* Genome-wide association study and genomic prediction using parental and breeding populations of Japanese pear (Pyrus pyrifolia Nakai). *Scientific Reports* **8**, 11994 (2018).

95. Su, H. *et al.* Leaf-level plasticity of Salix gordejevii in fixed dunes compared with lowlands in Hunshandake Sandland, north China. *Journal of Plant Research* **122**, 611–622 (2009).

96. Poorter, H. Plant Growth and Carbon Economy. *Encyclopedia of Life Sciences* (2002). doi:10.1038/npg.els.0003200

97. Li, X., Schmid, B., Wang, F. & Paine, C. E. T. Net assimilation rate determines the growth rates of 14 species of subtropical forest trees. *PLoS ONE* **11**, 1–13 (2016).

98. Fiorani, F. & Schurr, U. Future scenarios for plant phenotyping. *Annual Review of Plant Biology* **64**, 267–291 (2013).

99. Das Choudhury, S., Bashyam, S., Qiu, Y., Samal, A. & Awada, T. Holistic and component plant phenotyping using temporal image sequence. *Plant Methods* **14**, 1–21 (2018).

100. Beaulieu, J., Doerksen, T. K., Mackay, J., Rainville, A. & Bousquet, J. Genomic selection accuracies within and between environments and small breeding groups in white spruce. 1–16 (2014).

101. Tsai, H. Y. *et al.* Genomic prediction and GWAS of yield, quality and disease-related traits in spring barley and winter wheat. *Scientific Reports* **10**, 1–15 (2020).

102. Suontama, M. *et al.* Ef fi ciency of genomic prediction across two Eucalyptus nitens seed orchards with different selection histories. 370–379 (2019). doi:10.1038/s41437-018-0119-5

103. Kwong, Q. Bin *et al.* Evaluation of methods and marker Systems in Genomic Selection of oil palm (Elaeis guineensis Jacq.). *BMC Genetics* **18**, 1–9 (2017).

104. Ng, F. S. P. Age of trees in tropical rainforests estimated by timing of wood decay. *Journal of Tropical Forest Science* **25**, 437–441 (2013).

105. Fichtler, E., Clark, D. A. & Worbes, M. Age and Long-term Growth of Trees in an Old-growth Tropical Rain Forest, Based on Analyses of Tree Rings and 14C. *Biotropica* **35**, 306–317 (2003).

106. Resende, M. F. R. *et al.* Accelerating the domestication of trees using genomic selection: Accuracy of prediction models across ages and environments. *New Phytologist* **193**, 617–624 (2012).

107. Lado, B., Barrios, P. G., Quincke, M., Silva, P. & Gutiérrez, L. Modeling genotype × Environment

interaction for genomic selection with unbalanced data from a wheat breeding program. *Crop Science* **56**, 2165–2179 (2016).

108. Heffner, E. L., Sorrells, M. E. & Jannink, J. L. Genomic selection for crop improvement. *Crop Science* **49**, 1–12 (2009).

109. Wang, T., Hamann, A., Spittlehouse, D. L. & Murdock, T. Q. ClimateWNA-high-resolution spatial climate data for western North America. *Journal of Applied Meteorology and Climatology* **51**, 16–29 (2012).

110. Whitlock, C. H. *et al.* Release 3 NASA Surface Meteorology and Solar Energy Data Set for Renewable Energy Industry Use. *Proceedings of Rise and Shine* **1**, 1829–1841 (2000).

111. Saint Pierre, C. *et al.* Genomic prediction models for grain yield of spring bread wheat in diverse agro-ecological zones. *Scientific Reports* **6**, 1–11 (2016).

112. Thistlethwaite, F. R. *et al.* Genomic selection of juvenile height across a single-generational gap in Douglas-fir. *Heredity* **122**, 848–863 (2019).

113. Collevatti, R. G. *et al.* A genome-wide scan shows evidence for local adaptation in a widespread keystone Neotropical forest tree. *Heredity* **123**, 117–137 (2019).

114. Ortiz-Bibian, M. A. *et al.* Genetic Variation in Abies religiosa for Quantitative Traits and Delineation of Elevational and Climatic Zoning for Maintaining Monarch Butterfly Overwintering Sites in Mexico, considering Climatic Change. *Silvae Genetica* **66**, 14–23 (2017).

115. George, J. P. *et al.* Genetic variation, phenotypic stability, and repeatability of drought response in European larch throughout 50 years in a common garden experiment. *Tree Physiology* **37**, 33–46 (2017).

116. Givnish, T. Adaptation to Sun and Shade: A Whole-plant Perspective. *Aust. J. Plant Physiol.* 63–92 (1988).

117. Plomion, C. *et al.* Understanding the genetic bases of adaptation to soil water deficit in trees through the examination of water use efficiency and cavitation resistance: maritime pine as a case study. *Journal of Plant Hydraulics* **3**, 008 (2016).

118. Ren, C. *et al.* Responsiveness of soil nitrogen fractions and bacterial communities to afforestation in the Loess Hilly Region (LHR) of China. *Scientific Reports* **6**, 1–11 (2016).

119. Hewitt, N. *et al.* Taking stock of the assisted migration debate. *Biological Conservation* **144**, 2560–2572 (2011).

120. Sniezko, R. A. & Koch, J. Breeding trees resistant to insects and diseases: putting theory into application. *Biological Invasions* **19**, 3377–3400 (2017).

121. Zipperer, W., Sisinni, S., Pouyat, R. & Foresman, T. Urban tree cover: an ecological perspective. *Urban Ecosystems* **1**, 229–246 (1997).

122. Rowntree, R. A. Urban forest ecology: Conceptual points of departure. *Journal of Arboriculture* **24**, 62–71 (1998).

123. Parsons, S. E. & Frank, S. D. Urban tree pests and natural enemies respond to habitat at different spatial scales. *Journal of Urban Ecology* **5**, 1–15 (2019).

124. Jones, F. A. & Hubbell, S. P. Demographic spatial genetic structure of the Neotropical tree, Jacaranda copaia. *Molecular Ecology* **15**, 3205–3217 (2006).

125. Fajardo, A. & Siefert, A. Intraspecific trait variation and the leaf economics spectrum across resource gradients and levels of organization. *Ecology* **99**, 1024–1030 (2018).

126. Kuhn, D. N. *et al.* Genetic map of mango: A tool for mango breeding. *Frontiers in Plant Science* **8**, 1–11 (2017).

127. Urquía, D. *et al.* Psidium guajava in the Galapagos Islands: Population genetics and history of an invasive species. *bioRxiv* 1–21 (2018). doi:10.1101/402693

128. Bressler, A., Vidon, P., Hirsch, P. & Volk, T. Valuation of ecosystem services of commercial shrub willow (Salix spp.) woody biomass crops. *Environmental Monitoring and Assessment* **189**, (2017).

129. Lafleur, B., Labrecque, M., Arnold, A. A. & Bélanger, N. Organic carbon accumulation in topsoil following afforestation with willow: Emphasis on leaf litter decomposition and soil organic matter quality. *Forests* **6**, 769–793 (2015).

130. Mckown, A. D. *et al.* Geographical and environmental gradients shape phenotypic trait variation and genetic structure in Populus trichocarpa. *New Phytologist* **201**, 1263–1276 (2014).

131. Hay, E. H. A. *et al.* Genomic predictions combining SNP markers and copy number variations in Nellore cattle. *BMC Genomics* **19**, 1–8 (2018).

132. Sow, M. D. *et al.* Epigenetics in Forest Trees: State of the Art and Potential Implications for Breeding and Management in a Context of Climate Change. *Advances in Botanical Research* **88**, 387–453 (2018).

133. Amaral, J. *et al.* Advances and promises of epigenetics for forest trees. *Forests* **11**, 1–21 (2020).

134. de Villemereuil, P., Mouterde, M., Gaggiotti, O. E. & Till-Bottraud, I. Patterns of phenotypic plasticity and local adaptation in the wide elevation range of the alpine plant Arabis alpina. *Journal of*

Ecology **106**, 1952–1971 (2018).

135. Gugger, P. F., Fitz-Gibbon, S., Pellegrini, M. & Sork, V. L. Species-wide patterns of DNA methylation variation in Quercus lobata and their association with climate gradients. *Molecular Ecology* **25**, 1665–1680 (2016).

136. Le Gac, A. L. *et al.* Winter-dormant shoot apical meristem in poplar trees shows environmental epigenetic memory. *Journal of Experimental Botany* **69**, 4821–4837 (2018).

137. Liu, G. E. *et al.* Analysis of copy number variations among diverse cattle breeds. *Genome Research* **20**, 693–703 (2010).

138. Brenig, B. *et al.* Molecular genetics of coat colour variations in White Galloway and White Park cattle. *Animal Genetics* **44**, 450–453 (2013).

139. Libbrecht, M. W. & Noble, W. S. Machine learning applications in genetics and genomics. *Nature Reviews Genetics* **16**, 321–332 (2015).

140. Ingvarsson, P. Prospects and utilities of genomic selection in forest tree breeding A (very) short history of genomic selection.

141. Kumar, S. *et al.* Genomic selection for fruit quality traits in apple (Malus×domestica Borkh.). *PLoS ONE* **7**, 1–10 (2012).

142. Roth, M. *et al.* Genomic prediction of fruit texture and training population optimization towards the application of genomic selection in apple. *Horticulture Research* **7**, (2020).

143. Minamikawa, M. F. *et al.* Genome-wide association study and genomic prediction using parental and breeding populations of Japanese pear (Pyrus pyrifolia Nakai). *Scientific Reports* **8**, (2018).

144. Minamikawa, M. F. *et al.* Genome-wide association study and genomic prediction in citrus: Potential of genomics-assisted breeding for fruit quality traits. *Scientific Reports* **7**, 1–2 (2017).

145. Gois, I. B. *et al.* Genome wide selection in citrus breeding. *Genetics and Molecular Research* **15**, 1–14 (2016).

146. Imai, A. *et al.* Single-step genomic prediction of fruit-quality traits using phenotypic records of non-genotyped relatives in citrus. *PLoS ONE* **14**, 1–14 (2019).

147. Ferrão, L. F. V. *et al.* Accurate genomic prediction of Coffea canephora in multiple environments using whole-genome statistical models. *Heredity* **122**, 261–275 (2019).

148. McElroy, M. S. *et al.* Prediction of cacao (Theobroma cacao) resistance to moniliophthora spp. diseases via genome-wide association analysis and genomic selection. *Frontiers in Plant Science* **9**, 1–12 (2018).

149. Alberto Romero Navarro, J. *et al.* Application of genome wide association and genomic prediction for improvement of cacao productivity and resistance to black and frosty pod diseases. *Frontiers in Plant Science* **8**, (2017).

150. de Oliveira, E. J. *et al.* Genome-wide selection in cassava. *Euphytica* **187**, 263–276 (2012).

151. Wolfe, M. D. *et al.* Prospects for Genomic Selection in Cassava Breeding. *The Plant Genome* **0**, 0 (2017).

152. Kayondo, S. I. *et al.* Genome-wide association mapping and genomic prediction for CBSD resistance in Manihot esculenta. *Scientific Reports* **8**, 1–11 (2018).

153. Nyine, M. *et al.* Genomic Prediction in a Multiploid Crop: Genotype by Environment Interaction and Allele Dosage Effects on Predictive Ability in Banana. *The Plant Genome* **0**, 0 (2018).

154. Kwong, Q. Bin *et al.* Evaluation of methods and marker Systems in Genomic Selection of oil palm (Elaeis guineensis Jacq.). *BMC Genetics* **18**, (2017).

155. Kwong, Q. Bin *et al.* Genomic selection in commercial perennial crops: Applicability and improvement in oil palm (Elaeis guineensis Jacq.). *Scientific Reports* **7**, (2017).

156. Souza, L. M. *et al.* Genomic Selection in Rubber Tree Breeding: A Comparison of Models and Methods for Managing G×E Interactions. *Frontiers in Plant Science* **10**, 1–14 (2019).

157. Cros, D. *et al.* Within-family genomic selection in rubber tree (Hevea brasiliensis) increases genetic gain for rubber production. *Industrial Crops and Products* **138**, 111464 (2019).

158. De Azevedo Peixoto, L., Laviola, B. G., Alves, A. A., Rosado, T. B. & Bhering, L. L. Breeding Jatropha curcas by genomic selection: A pilot assessment of the accuracy of predictive models. *PLoS ONE* **12**, 1–16 (2017).

159. Westbrook, J. W. *et al.* Optimizing genomic selection for blight resistance in American chestnut backcross populations: A trade-off with American chestnut ancestry implies resistance is polygenic. *Evolutionary Applications* **13**, 31–47 (2020).

160. Nagano, S. *et al.* SNP genotyping with target amplicon sequencing using a multiplexed primer panel and its application to genomic prediction in Japanese Cedar, Cryptomeria japonica (L.f.) D.Don. *Forests* **11**, (2020).

161. Kainer, D., Stone, E. A., Padovan, A., Foley, W. J. & Külheim, C. Accuracy of genomic prediction for foliar terpene traits in Eucalyptus polybractea. *G3: Genes, Genomes, Genetics* **8**, 2573–2583 (2018).

162. Ballesta, P., Bush, D., Silva, F. F. & Mora, F. Genomic predictions using low-density SNP markers, pedigree and GWAS information: A case study with the non-model species Eucalyptus cladocalyx. *Plants* **9**, (2020).

163. Klápště, J. *et al.* Marker Selection in Multivariate Genomic Prediction Improves Accuracy of Low Heritability Traits. *Frontiers in Genetics* **11**, 1–15 (2020).

164. Ballesta, P., Serra, N., Guerra, F. P., Hasbún, R. & Mora, F. Genomic prediction of growth and stem quality traits in Eucalyptus globulus Labill. at its southernmost distribution limit in Chile. *Forests* **9**, 1–18 (2018).

165. Ballesta, P., Maldonado, C., Pérez-Rodríguez, P. & Mora, F. SNP and haplotype-based genomic selection of quantitative traits in Eucalyptus globulus. *Plants* **8**, 1–18 (2019).

166. Mphahlele, M. M. *et al.* Expected benefits of genomic selection for growth and wood quality traits in Eucalyptus grandis. *Tree Genetics and Genomes* **16**, (2020).

167. Tan, B., Grattapaglia, D., Wu, H. X. & Ingvarsson, P. K. Genomic relationships reveal significant dominance effects for growth in hybrid Eucalyptus. *Plant Science* **267**, 84–93 (2018).

168. A, K. Genomic selection and g enome-wide association study in Populus trichocarpa and Pinus taeda. (2016).

169. Ukrainetz, N. K. & Mansfield, S. D. Assessing the sensitivities of genomic selection for growth and wood quality traits in lodgepole pine using Bayesian models. *Tree Genetics and Genomes* **16**, (2020).

170. Resende, J. F. R. *et al.* Accuracy of genomic selection methods in a standard data set of loblolly pine (Pinus taeda L.). *Genetics* **190**, 1503–1510 (2012).

171. Zapata-Valenzuela, J. *et al.* SNP markers trace familial linkages in a cloned population of Pinus taeda-prospects for genomic selection. *Tree Genetics and Genomes* **8**, 1307–1318 (2012).

172. Zapata-Valenzuela, J., Whetten, R. W., Neale, D., McKeand, S. & Isik, F. Genomic estimated breeding values using genomic relationship matrices in a cloned population of loblolly pine. *G3: Genes, Genomes, Genetics* **3**, 909–916 (2013).

173. Munoz, P. R. *et al.* Genomic relationship matrix for correcting pedigree errors in breeding populations: Impact on genetic parameters and genomic selection accuracy. *Crop Science* **54**, 1115–1123 (2014).

174. Isik, F. *et al.* Plant Science Genomic selection in maritime pine. *Plant Science* **242**, 108–119 (2016).

175. Bartholomé, J. *et al.* Performance of genomic prediction within and across generations in maritime pine. *BMC Genomics* **17**, 1–14 (2016).

176. Li, Y. *et al.* Genomic selection for non-key traits in radiata pine when the documented pedigree is corrected using DNA marker information. *BMC Genomics* **20**, 1–10 (2019).

177. Calleja-Rodriguez, A. *et al.* Evaluation of the efficiency of genomic versus pedigree predictions for growth and wood quality traits in Scots pine. *BMC Genomics* **21**, 1–17 (2020).

178. Chen, Z. Q. *et al.* Accuracy of genomic selection for growth and wood quality traits in two control-pollinated progeny trials using exome capture as the genotyping platform in Norway spruce. *BMC Genomics* **19**, (2018).

179. Lenz, P. R. N. *et al.* Genomic prediction for hastening and improving efficiency of forward selection in conifer polycross mating designs: an example from white spruce. *Heredity* **124**, 562–578 (2020).

180. Fuentes-Utrilla, P. *et al.* QTL analysis and genomic selection using RADseq derived markers in Sitka spruce: the potential utility of within family data. *Tree Genetics and Genomes* **13**, (2017).

181. Thistlethwaite, F. R. *et al.* Genomic prediction accuracies in space and time for height and wood density of Douglas-fir using exome capture as the genotyping platform. *BMC Genomics* **18**, 1–16 (2017).

182. Sawitri *et al.* Potential of Genome-Wide association studies and Genomic Selection to improve productivity and quality of commercial timber species in tropical rainforest, a case study of Shorea platyclados. *Forests* **11**, (2020).

CAPÍTULO IV

Evaluación de la precisión de la predicción genómica para el manejo y conservación de poblaciones aisladas de árboles naturales

Tree Genetics & Genomes (2021) 17:12
https://doi.org/10.1007/s11295-020-01489-1

ORIGINAL ARTICLE

Evaluating the accuracy of genomic prediction for the management and conservation of relictual natural tree populations

Sebastián Arenas[1,2] · Andrés J. Cortés[3,4] · Alicia Mastretta-Yanes[5,6] · Juan Pablo Jaramillo-Correa[2]

Received: 27 September 2020 / Revised: 16 December 2020 / Accepted: 30 December 2020

Abstract

Studying and understanding the evolution of relictual natural populations is critical for developing conservation initiatives of endangered species, such as management in situ and assisted migration. Recently, genomic and bioinformatics tools have promised a wide avenue for developing more efficient programs. Genomic prediction (GP) models are one of such tools; although, in trees, only some successful examples exit. They have mostly been used to increase predictive ability in commercial traits and reduce breeding cycle length. Thus, it remains to be tested whether GP can be extended for the management and conservation of natural small and secluded populations. Here, we explored such a possibility in a pilot study to predict the performance of introduced saplings in a managed population of sacred fir (*Abies religiosa*; Pinaceae) in central Mexico. We genotyped over 200 naturally re-generated and introduced individuals with 2286 single nucleotide polymorphisms (SNP), derived from genotyping by sequencing, and used them to develop GP models for growth and physiological traits. After testing different training and validation datasets, and determining predictive ability of "across-groups" models with cross-validation techniques, acceptable predictive abilities (r_y) were obtained for growth during the previous growing season, water potential, stem diameter, and aboveground biomass (0.36, 0.27, 0.26, and 0.24, respectively). The best models were always those built with natural saplings and used to predict the early performance of introduced individuals in the same environment, although fair predictabilities were also obtained when predicting performance between natural populations. Model fine-tuning resulted in reduced datasets of approximately 700 SNPs that helped optimizing phenotype predictability, particularly for water potential, for which r_y was up to 0.28. These pilot-scale results are preliminary but encouraging and justify additional research efforts for implementing GP in small and secluded natural populations, particularly for endangered non-model species.

Keywords *Abies* · Forest management · Genomic prediction (GP) · Mexico · Predictive ability · Secluded natural populations

Introduction

Understanding how long-lived species evolve in small secluded population has become an integral part of adaptive forest management and conservation, especially for endangered species that persist only in restricted geographical ranges (Pautasso 2009; Bezemer et al. 2019). This is required for forecasting the effects of climate change on range shift, local adaptation, and for predicting individuals' suitability for reforestation (e.g., Wang et al. 2006; Le et al. 2012). Studies on

Communicated by J. Beaulieu

✉ Juan Pablo Jaramillo-Correa
jaramillo@ecologia.unam.mx

[1] Posgrado en Ciencias Biológicas, Universidad Nacional Autónoma de México, Ciudad de Mexico, Mexico

[2] Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad de Mexico, Mexico

[3] Corporación Colombiana de Investigación Agropecuaria (AGROSAVIA) – CI La Selva, Rionegro, Colombia

[4] Facultad de Ciencias Agrarias – Departamento de Ciencias Forestales, Universidad Nacional de Colombia – Sede Medellín, Medellin, Colombia

[5] Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), Liga Periférico Insurgentes Sur 4903, Col. Parques del Pedregal, Tlalpan, Ciudad de Mexico, Mexico

[6] Consejo Nacional de Ciencia y Tecnología, Benito Juárez (CONACYT), CDMX, México, Avenida Insurgentes Sur 1582, Crédito Constructor, Benito Juárez, Ciudad de Mexico, Mexico

Published online: 03 February 2021

population divergence conducted in tree populations have traditionally used modest sample sizes and mainly focused on estimating genetic diversity and heritability for commercially important traits through Q_{ST} and F_{ST} statistics (Holliday et al. 2016; Gentili et al. 2018). These studies have revealed certain trends in small populations (e.g., low genetic diversity, increased relatedness, weak within-population genetic structure, reduced incoming gene flow, and increased inbreeding and linkage disequilibrium "LD"); they may all affect local adaptation and compromise species survival (Finlay et al. 2017; Bezemer et al. 2019; Garot et al. 2019). Therefore, considering these trends and their evolutionary consequences is key for promoting effective conservation and management in relictual populations (Cobo-Simón et al. 2020).

Transferring tree germplasm has shaped the management, ecology, and genetic diversity of forests, both natural and planted (introduced) (Aitken and Whitlock 2013; Lin et al. 2018); recently, it has arisen as an important strategy for assisting populations to cope with climate change (Loo et al. 2014; Sáenz-Romero et al. 2016). Transferred germplasm has further been used to grow trees for numerous purposes; ranging from the production of wood and non-wood products, to the provision of ecosystem services, such as forest restoration and ecological mitigation (Le et al. 2012; Dawson et al. 2014). However, plantations based on transferred germplasm also raise concerns and difficulties, such as the possibility of outbreeding depression, which can lead to significant bottlenecks (Koskela et al. 2014). Indeed, when incorporated in assisted migration programs, transfer of individuals may affect species' prospects in a changing climate, especially if they do not encompass adaptive genetic variation to local environments (Aitken et al. 2008; Aitken and Bemmels 2015; George et al. 2017). From this point of view, it would be interesting to explore whether predictive tools can be developed to pre-select locally adapted individuals for reforestation. Such tools may also provide support for assisted migration plans and improving transfer success and long-term viability.

The development of modern genomic tools and the continuous decrease of sequencing costs (Nystedt et al. 2013; Thomson 2014) has allowed geneticist to integrate information from large SNP datasets into more accurate predictive models of phenotypic variation through GP models (Ingvarsson et al. 2016; Suontama et al. 2019). These sets have revolutionized plant genetic programs, which may now incorporate genetic variants driving complex phenotypes (Grattapaglia 2014; Minamikawa et al. 2018). GP has also helped to overcome the limitations of marker-based approaches, which traditionally search and map a small amount of quantitative trait loci (QTLs) with large phenotypic effects (Meuwissen et al. 2001; Barton et al. 2017). Although selective breeding is still limited in forest trees, growing genomic resources (Isik 2014; Grattapaglia et al. 2018) has prompted various promising case studies (Beaulieu et al. 2014a; Isik et al. 2016; Müller et al. 2017; Li et al. 2019), even some in species that are still in early progeny trial stages (Rambolarimanana et al. 2018). Such cases somehow lead to the question of whether predictive models can also be built for natural tree populations. Unfortunately, initial attempts in unrelated families failed to build accurate models (e.g., Beaulieu et al. 2014b), although some predictive ability was obtained when training and validating models within the same population (Beaulieu et al. 2014a; Suontama et al. 2019), which likely included local relatedness and provenance structure (Ratcliffe et al. 2017).

For breeding populations of forest trees with high relatedness, a few thousands of markers distributed throughout the genome have been shown to be enough for having a fair predictive ability of complex phenotypes, such as growth and wood properties (Müller et al. 2017; Lenz et al. 2020a; Klápště et al. 2020; Thistlethwaite et al. 2020). Such results hint that GP models could indeed be developed for natural populations if they have been historically small, that is, populations where there is high genetic relatedness because of recurrent consanguinity over various generations. Indeed, individuals evolving in small populations for long periods of time may share a high proportion of alleles (typically measured as relationship due to coancestry), which could result in increased LD between genotyped markers and the genetic variants underlying phenotypic variation (Campoy et al. 2016). In other words, if LD is extensive as a result of increased consanguinity (Caldwell et al. 2006; Wang et al. 2016), and if natural selection has been strong enough over the years, we hypothesize that models can be built to predict phenotypes in such small secluded natural populations for estimating a sizable proportion of genomic h^2.

Difficulties for developing such models in natural stands include the absence of obvious (and controlled) pedigree relationships within the population, a high diversity of age classes among individuals, and the possibility that parental trees differ between consecutive years. Besides, GP studies are preferentially performed with replicated material in multi-environment trials, which allows estimating G × E interactions (Gamal El-Dien et al. 2015; Ratcliffe et al. 2019; Thistlethwaite et al. 2019) and including environmental-specific marker effects (Resende et al. 2012a; Suontama et al. 2019). Such a possibility would be largely limited (if not impossible) for natural populations, which would likely result in limited model transferability. However, such models may still help defining more reliable management and conservation strategies.

Here, we performed a pilot study within a reforestation test trial in Central Mexico for exploring the feasibility of GP modeling in small and secluded natural populations. The study site was simultaneously repopulated by naturally regenerated and introduced sacred fir (*Abies religiosa* (Kunth) Schltdl. & Cham) individuals after selective wood extraction in 2004. Sacred fir is a tropical conifer distributed in relatively small and isolated populations scattered along the Trans-Mexican Volcanic Belt (central Mexico) at high elevations (2800–3500 m.a.s.l.; Méndez-González et al. 2017). Such a distribution makes it especially

vulnerable to increased temperatures and drought resulting from climate change, given that migration rates could be insufficient to track down such changes (Sáenz-Romero et al. 2012). This species has shown significant heritable phenotypic variation associated to climate (Ortíz-Bibian et al. 2017; Cruz-Nicolás et al. 2020) and marked differences in response to elevation gradients (Sáenz-Romero et al. 2012; Carbajal-Navarro et al. 2019). Assisted migration appears thus a viable conservation solution for sacred fir (Sáenz-Romero et al. 2016), although initial attempts have had mixed results (Carbajal-Navarro et al. 2019). Selecting pre-adapted individuals may help such management and conservation plans, but one first has to demonstrate that predictive models actually work in wild populations.

We examined whether GP models could be built for this small pilot stand by considering phenotypic traits related to sapling performance and survival, which are at the base of any long-term plantation/population conservation program. Specifically, we first estimated heritability for ten traits and determined the predictive ability of genetic estimates of breeding values (GEBVs). We then compared such predictive ability across different training groups to identify the ideal dataset composition for GP modeling. To do so, we aimed testing whether the early performance of introduced individuals could be predicted from models constructed for naturally regenerated local plants, and determine the transferability of predictive models between sites, in spite of our inability to accurately include environment-specific factors (e.g., genotype × environment interactions). Although limited because of sample size and marker density, results from this exploratory study suggest that GP modeling is feasible in historically small and secluded populations and call for additional research efforts at a much larger scale.

Materials and methods

Study site and sampling

The study sites are within a managed forest area in the buffer zone of the Nevado de Toluca flora and fauna protection area (NTFFPA), in central Mexico. Management focuses in the sustainable use of timber forest resources from different "ejidos" (communally owned and managed lands) with the assistance of Consejo Civil Mexicano para la Silvicultura Sostenible (CCMSS). Our sampling was carried out in two forest trials, Rincón de Guadalupe (19° 15′ N; 99° 57′ W) and San Bartolo (19° 14′ N; 99° 54′ W). These two sites are separated by 7.4 km and differ in elevation (2845 vs. 3000 m.a.s.l.), light intensity at noon (which respectively ranged from 292 to 487 μmol/m^{2}s and 371 to 1090 μmol/m^{2}s according to measures performed with a Digital Lux Meter (Model LX-101, Taiwan)), and several edaphic characteristics (Table 1). Selective wood extraction at both sites has generated forest clearings that have been repopulated by natural regeneration in San Bartolo (SB) and by both locally regenerated (NR) and introduced (RF) individuals in Rincón de Guadalupe. According to CCMSS, germplasm of RF individuals comes from nearby localities, although there are not records of the exact provenance. We sampled and measured 213 saplings from these three groups (81 SB-individuals, 54 NR-individuals and 78 RF-individuals) (Table S1). All plants were devoid of apparent biotic attacks and were at least 3 m apart from other trees. They were all between 1.2 and 2 m in height and had well-formed stems, which makes sapling performance and survival estimation reliable. We inferred individual-age from bud-scar counts as in Hankin et al. (2018) and Urza and Sibold (2013). Ages ranged from 6 to 12 years, and trees were classified in seven age classes (i.e., 6, 7, 8, 9, 10, 11, and 12 year-old individuals; Table S1) to account for age effects in analyses below.

Samples for genomic analyses consisted of 2-year-old needles. Phenotypic measures were taken for growth and physiological traits including stem diameter at 20 cm from ground level (SD, cm), total height (TH, cm), aboveground biomass (AGB, Kg), crown ratio (CR, cm), growth during the 2016 growing season (G_{2016}, cm), mean wood mean density (WD, g/cm^3), water potential (Ψ, MPa), needle relative water content (RWC, %), specific foliar area (SLA, cm^2/g), and relative growth rate (RGR, g/g day).

For estimating AGB, we used the general allometric equations described by Chojnacky et al. (2014) for *Abies* individuals with diameters below 69 cm. For four traits (WD, Ψ, RWC and SLA), vegetative tissue was sampled in the upper half of the crown, immediately placed on ice, and stored at − 10 °C until further processing in the laboratory (Lenz et al. 2017). Ψ was measured using a Schölander pressure chamber following Cochard et al. (1992) and Larbi et al. (2003). Because trees within a flora and fauna protection area cannot be felled at young ages, WD was determined for each plant by averaging values determined for two pieces of wood collected from main branches with base diameters between 0.5 and 0.7 cm. Because branches usually have high proportions of compression wood (Livingston et al. 2004), these values should only be viewed as *proxies*. Cross-sectional discs were cut from each branch at a distance of 5 cm from the base (Dibdiakova and Vadla 2012). We then measured the water displacement of each wood piece within a graduated test tube to determine fresh volume. Afterwards, dry mass was determined with an electronic balance (precision of 0.001 g) after drying in an oven at 75 °C for 24 h. WD was estimated following Williamson and Wiemann (2010) as follows (Eq. 1):

$$\mathrm{WD} = \frac{\text{dry mass (g)}}{\text{fresh volume (cm}^3\text{)}} \quad (1)$$

RWC was measured by duplicate using 20 randomly selected needles per plant, which were weighed using an

Table 1 Edaphic variables determined at 0–15 cm depths in soils around sacred fir (*Abies religiosa*) individuals from two test sites in central Mexico. Means followed by different letters are significantly different from each other ($p < 0.05$). SE, standard error; Ns, not significant

Edaphic characteristic	San Bartolo			Rincón de Guadalupe		
	Mean		SE	Mean		SE
pH (H_2O)	5.5^{a}	±	0.30	6.1^{b}	±	0.10
Exchangeable cations (cmol/kg)	258.4^{a}	±	44.00	128.3^{b}	±	9.00
Total C (%)	16.0^{Ns}	±	2.4	15.4^{Ns}	±	2.3
Total N (%)	1.71^{a}	±	0.1	1.54^{b}	±	0.1

electronic balance; estimations were performed following Soltys-Kalina et al. (2016) (Eq. 2):

$$\mathrm{RWC} = \frac{\text{fresh mass (g)} - \text{dry mass (g)}}{\text{turgid mass (g)} - \text{dry mass (g)}} \times 100 \quad (2)$$

SLA (cm^2/g), defined as the radio fresh leaf area/dry mass (Poorter 2002; Liu et al. 2017), was estimated from the same needles above using images obtained with a scanner (HP Scanjet G3110). Leaf area was calculated using ImageJ v.1.8 (Abramoff et al. 2004).

Relative growth rate (RGR) was estimated as follows (Eq. 3):

$$\mathrm{RGR} = \frac{\mathrm{Ln}(\mathrm{AGB}_{2017}) - \mathrm{Ln}(\mathrm{AGB}_{2016})}{t_{2017} - t_{2016}} \quad (3)$$

where AGB_{2016} and AGB_{2017} are respectively the plant dry aboveground biomass for the previous growing season (i.e., 2016) and for the year during which samples were taken (2017); $t_{2017} - t_{2016}$ denotes the time spanned (in numbers of days) between the end of the previous growing season and the sampling date (Hoffmann and Poorter 2002).

Differences between groups of individuals (i.e., between origins and sites) were determined for each trait through one-way analyses of variance (ANOVA), followed by Tukey tests of honesty (HSD). Correlations were performed between phenotypic traits in the *PerformanceAnalytics* package (Brian et al. 2019). All analyses were carried out in R v.3.4.4 (The R core Team 2019).

Sequencing, assembly, and genotyping

Total genomic DNA was extracted from needles following the protocol described by Telfer et al. (2013) and using the DNAeasy Plant Mini Kit (Qiagen, Germany). Library preparation was performed according to Poland et al. (2012) after DNA restriction with *PstI*. Single-end sequence reads (80 bp) were obtained from an Illumina HiSeq 2500 sequencer at the Institute of Integrative Biology and Systems at Université Laval (http://www.ibis.ulaval.ca/en/services-2/genomic-analysis-platform/). Demultiplexing and quality filtering, de novo assembly, read alignment, and SNP calling were performed with an *Ipyrad* v0.7.23 (Eaton 2014) pipeline. Assembling parameters included a clustering threshold of 0.9, a mindepth of 8, and a maximum barcode mismatch of 0. Each sampling site had to be represented in at least one individual for a SNP to be called. Subsequently, monomorphic reads, variants with missing call rates above 30%, and samples with minimum allele frequencies (MAF) of 5% and in Hardy-Weinberg disequilibrium (p value $< 1 \times 10^{-6}$) (Minamikawa et al. 2018) were eliminated with Plink v1.07 (Purcell 2010). Missing genotypic data were imputed with TASSEL 5 (Bradbury et al. 2007) using LD *K*-nearest neighbor (Money et al. 2015). In total, 12.7% of missing genotype data were imputed. Furthermore, genetic structure was estimated among ages-classes according to Chung et al. (2003), for evaluating its influence on genetic differentiation, which could bias GP models.

Analysis of relationship, population structure, and LD among groups of individuals

The resulting panel of 2286 SNPs was successfully genotyped for 201 samples distributed among the three groups of individuals (i.e., 77 for RF, 51 for NR and 73 for SB), and used to estimate a realized genomic relationship matrix among individuals (G) with the "Gmatrix" function in the AGHmatrix package (Amadeu et al. 2016); we used options described by VanRaden (2008). To test for population structure, this G matrix was spectrally decomposed with the function *eigen()* in R v.3.4.4 (The R core Team 2019). Such a spectral decomposition allowed estimating the variance captured by the first eigenvector, and thus the extent of population structure among groups of individuals. Following Marco de Lima et al. (2019), we inferred the relationship structure of our sample by estimating paired-relatedness values between individuals from our SNP data, using threshold values for identifying unrelated (0.00), half-cousins (0.0625), first-cousins (0.125), half-sibs (0.25), and full-sibs (0.50), respectively. Population structure was further evaluated with a Principal Component Analysis (PCA) with package *SNPRelate* (Zheng et al. 2012), and through a pairwise F_{ST} matrix (Weir and Cockerham 1984) calculated with the *genet.dist* function in *hierfstat* (Goudet and Jombart 2015). To evaluate the contribution of LD on the

accuracy and efficiency of GP, LD between pairs of SNPs was estimated with a pairwise correlation coefficient (r^2) with TASSEL 5 (Bradbury and Zhang 2007). Then, as another way to determine if a wide enough spectrum of genetic relatedness among individuals was retained (given our markers density), we used mean LD values to estimate the effective population size (N_e) for each group based on the LDN_e method implemented in NeEstimator v2.01 (Do et al. 2014). Following Müller et al. (2017), and given our relatedness matrix (see Results), we assumed a random mating model within each group. Confidence intervals for each estimate were obtained with a parametric method in which the number of independent alleles was used as the number of degrees of freedom in a *chi*-square distribution (Do et al. 2014). Furthermore, following May et al. (2013), we evaluated differences in the average and distribution of the minor allele frequencies (MAF) among age classes.

Genomic prediction analyses

Performance of genomic prediction (GP) models and variance components for each trait were obtained using the *BGLR* (Bayesian generalized linear regression) package (Pérez-Rodríguez and de los Campos 2014) for R v.3.4.4 (The R core Team 2019). As environmental effects are of major importance in forestry (Cappa et al. 2016; Thistlethwaite et al. 2017), and to account for in situ differences (environment) and age classes (described above), genetic parameters for all phenotypic growth and physiology traits were estimated following Bouvet et al. (2020), by using a mixed linear model (also called "animal model"; Lenz et al. 2020b) fitted as follows:

$$y = Xb + Za + e \tag{4}$$

where y is a vector of a phenotypic trait; a is a normally distributed vector, $a \sim (0, G\ \sigma_a{}^2)$, of individual random additive genetic effects (with G being the genomic relationships matrix among individuals calculated above and $\sigma_a{}^2$ the additive genetic variance (Müller et al. 2017), b is a vector of fixed effects (including the general mean, age class and experimental site effects), and e is a vector representing the random residual effects, which were normally distributed: $e \sim N(0, I\ \sigma_e{}^2)$, where I is an identity matrix and $\sigma_e{}^2$ the residual variance. X and Z in Eq. (4) are the corresponding incidence matrices for fixed effects and additive genetic (Chen et al. 2018; Zhou et al. 2020).

SNP effects were predicted for each individual trait using two different GP approaches applied on the above mixed linear model (4), the Bayesian Regression Ridge (BRR), and the reproducing kernel Hilbert space methods (RKHS). These allowed comparing prediction ability, pinpointing the best-predicting method, and building subsequent "across-groups" models.

The BRR was fitted assuming that vector a in Eq. (4) has a multivariate normal prior distribution with a variance common to all effects (Pérez-Rodríguez et al. 2010). Given that assuming such a distribution for natural populations could be problematic, we compared these results with those from RKHS (Gianola et al. 2006). This method is a semiparametric approach that allows inferring a given SNP marker function without making strong prior assumptions about the distribution of between-marker effects. Instead, this SNP marker function predicts genomic-enhanced genotypic values if the Gaussian Kernel encodes additive effects, which depends on a bandwidth parameter (h) (Cuevas et al. 2016). These two GP methods were then extended using a Gibbs sampling for estimating variance components. The number of iterations was set to 180,000, after discarding the 50,000 initial steps as burn-in, and using a thinning interval of 1000. Convergence of posterior distributions was verified using trace plots. Flat priors were used for all models.

Predictive ability and heritability of the models

Models were evaluated based on predictive ability (r_y), which was estimated per trait, for each group of individuals (i.e., SB, NR and RF) and for the whole sample set (combined model) as the Pearson's correlation coefficient between the vector of the observed phenotypic variable (y) and the genomic-estimated breeding values from cross validations (see below), i.e., $r(y, GEBV)$ of Eq. (4) (Müller et al. 2017; Bouvet et al. 2020). Then, models for Eq. (4) for each trait and group of individuals were used to calculate additive (σ_a^2) and residual (σ_e^2) variances (Sousa et al. 2019; Zhang et al. 2017), for estimating narrow sense heritability (h^2), an equivalent to the genomic heritability h_g^2, as in de los Campos et al. (2015):

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2} \tag{5}$$

Standard errors (SE) were calculated by assuming statistical independence between observations as follows:

$$SE = \frac{\sigma}{\sqrt{n}} \tag{6}$$

where σ is the standard deviation of r_y and n is the number of replicates (Gamal El-Dien et al. 2016).

We performed a cross-validation ("CV") by implementing a random subsampling partitioning of the data in two subsets.

The first one consisted of 90% of individuals, and was used as a training set (TRN) for building the model and estimating marker effects. The second subset (the remaining 10% of individuals) was used as testing set (TST) for predicting phenotypes based on marker effects. Such subsets were created for each trait evaluated within each group of individuals and for the whole sample. As a control, a 20-fold cross-validation was carried out following methodologies from Tan et al. (2017), with the same proportion of individuals used in the split sets, and using random allocations of individuals to the TRN and TST sets. Such a number of repetitions (instead of the more traditional 10-fold cross-validation) allowed for a lower probability of having no predictions for a given individual (at the expense of a higher probability of multiple predictions for another given tree). This process was repeated for each trait and method, except for evaluating the different TRN and TST sizes and compositions on the across-groups models (see below). For each replicate (i.e., fold), h^2 and r_y vales were retained for subsequent analyses. Finally, following Tan et al. (2017), significant differences for r_y and h^2 were evaluated among GP-methods (i.e., between BRR and RHKS), type of traits (i.e., between growth and physiological traits), and groups of individuals (i.e., between origins and sites) through ANOVAs; these were followed by Tukey tests of honesty (HSD) in the *PerformanceAnalytics* (Brian et al. 2019) package.

Impact of training group in prediction

To assess GP prediction ability for across-groups models, we adapted different CV techniques and changed both the TRN composition and size (Tan et al. 2017; Norman et al. 2018) and the TRN/TST ratio, at both the site and origin levels (using the BRR method only, see results). Initially, we tested the effect of TRN and TST composition by using four predictive models with varying TRN/TST ratios developed between group pairs, i.e., training group (G_{TRN}) and testing group (G_{TST}) (Resende et al. 2012a). In the first model, we aimed testing for phenotype prediction in reforested plants when using models constructed from naturally regenerated plants from the same site (i.e., NR was used as G_{TRN} and RF as G_{TST}; NR → RF). In a second time, and given that G × E plays an important evolutionary role (i.e., local adaptation; Sork 2018), we tested for model transferability, and used models constructed for regenerated plants at one site to predict the performance of plants reforested individuals at the other site (SB → RF). Similarly, the third (NR → SB) and fourth (SB → NR) models aimed predicting phenotypes across naturally regenerated groups; they were trained with data from one site and validated with data from the second site.

We then evaluated how changing TRN and TST size and composition affected the precision of the models above. To do so, we changed the relative size of each set in five different proportions of TRN/TST (while taking into account the number of individuals used for each model: $n = 128$ for NR → RF, $n = 150$ of SB → RF, and $n = 124$ for both NR → SB and SB → NR), so that TRN sets included all individuals from the training group (G_{TRN}) together with 0%, 10%, 20%, 30%, and 40% of individuals (selected randomly) from the testing group (G_{TST}). The rest of individuals from the testing groups were used for the TST set. The exact sizes used for each model are shown in Table S2. For each trait and model, 20 replicates were independently performed. Prediction abilities were computed for each model-trait combination as previously described for the CV partitions and testing subsets (Tan et al. 2017; Calleja-Rodriguez et al. 2020). These correlations were further used to select "top traits": those for which predictive ability increased as TRN size increased (from 0 to 40% of G_{TST}).

Most predictive SNP datasets for top traits

Given that complex models built with several thousands of markers might be difficult to implement in applied programs, we determined whether we could develop optimized models with a minimal number of predictive SNPs without significantly affecting predictive ability (Müller et al. 2017; Gutierrez et al. 2018). We built fourteen subsets of markers composed of 50, 100, 200, 300, 500, 600, 800, 1 K, 1.2 K, 1.3 K, 1.5 K, 1.75 K, 2 K, and 2.2 K SNPs for each selected "top trait" and compared r_y with those for the full 2286 SNP datasets. We performed 20 replicates for each subset. Each subset subsequently included the SNPs with the largest positive effects in the previous and smaller subset, as estimated from the retained GP models (Norman et al. 2018). Predictive abilities for this approach were fit by a curve. The minimum marker sets needed to recover "optimum" r_y values were compared across top traits using the *UpSetR* package (Lex et al. 2014) for R (The R core Team 2019), which aimed capturing the minimum set of highly predictive SNPs for each trait. Plots were drawn using the *ggplot2* package (Wickham 2010) for R (The R core Team 2019).

Results

Phenotypic differences among groups

There were significant differences at eight out of ten phenotypic traits (HT, SD, AGB, CR, G_{2016}, WD, Ψ, and RWC) between the naturally regenerated individuals from different sites (NR and SB) and between naturally

Table 2 Mean and standard errors (SE) for 10 phenotypic traits measured in sacred fir (*Abies religiosa*) individuals at two trial sites in central Mexico, followed by ANOVA results for testing for differences among groups: natural regeneration at both sites (SB and NR), and introduced saplings at Rincón de Guadalupe (RF). Different letters between groups denote significant mean differences ($p < 0.05$) according to Tukey's HSD

Phenotypic trait	San Bartolo			Rincon de Guadalupe						F	p
	Natural regeneration (SB)			Natural regeneration (NR)			Reforestation (RF)				
	Mean		SE	Mean		SE	Mean		SE		
Growth traits											
Total height (TH, cm)	121.10^{a}	±	3.60	140.70^{b}	±	6.50	93.60^{c}	±	3.20	29.40	***
Stem diameter (SD, cm)	1.82^{a}	±	0.07	1.61^{b}	±	0.08	0.96^{c}	±	0.03	52.80	***
Aboveground biomass (AGB, kg)	0.84^{a}	±	0.09	0.59^{b}	±	0.06	0.20^{c}	±	0.01	24.40	***
Crown ratio (CR, cm)	43.30^{a}	±	1.84	53.01^{b}	±	2.80	31.10^{c}	±	0.95	34.50	***
2016's growth (G_{2016}, cm)	25.80^{a}	±	1.80	26.60^{a}	±	2.02	10.50^{b}	±	1.20	31.70	***
Wood density (WD, g/cm^3)	0.51^{a}	±	0.01	0.54^{a}	±	0.02	0.41^{b}	±	0.01	28.20	***
Physiological traits											
Water potential (Ψ, MPa)	-0.31^{a}	±	0.01	-0.32^{a}	±	0.02	-0.47^{b}	±	0.02	27.60	***
Relative water content (RWC, %)	71.70^{a}	±	1.40	74.60^{a}	±	1.90	65.60^{b}	±	1.40	8.90	***
Specific leaf area (SLA, cm^2/g)	0.04^{a}	±	0.00	0.04^{a}	±	0.00	0.04^{a}	±	0.00	3.50	Ns
Relative growth rate (RGR, g/g/day)	1.43^{a}	±	0.13	1.31^{a}	±	0.11	1.53^{a}	±	0.15	0.70	Ns

Ns non-significant

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

regenerated and introduced saplings within the same site (NR and RF; ANOVA, $p < 0.05$). In general, and according to Tukey's tests, naturally regenerated individuals had higher means than the reforested ones (HSD, $p < 0.05$; Table 2), while plants from NR and SB groups only presented significant differences at certain growth traits. NR plants had higher means than SB individuals for two of these traits (140.7 ± 6.5 for TH and 53.01 ± 2.8 for CR), while SD and AGB were higher for saplings from SB than for NR plants (1.82 ± 0.07 for SD and 0.84 ± 0.09 for AGB). Some of these growth traits (e.g., SD, TH, and CR) were significantly correlated to each other (Fig. S1; $p < 0.05$).

SNP calling and genotyping

We obtained, in average, 2,803,267 raw reads per individual. After quality filtering, 31,462 consensus reads were retained and assembled de novo for identifying 373,267 SNPs. Many polymorphisms were however discarded because of low coverage, excessive missing data (higher than 30%), and minimum allelic frequencies below 5%. The final dataset consisted of 2286 SNPs (0.62% of the total) genotyped for 201 individuals evenly distributed across sites and origins. These were used to estimate individual relationships, population structure, and linkage disequilibrium and elaborate genomic prediction models.

Population structure, age effect, genomic relationship, and linkage disequilibrium assessment

The first four PCs of a PCA based on SNP genotypes explained 8.19% of the total genetic variation. Population structure was weak, with low segregation, and individuals did not show any obvious clustering, although naturally regenerated and introduced plants from Rincón de Guadalupe (NR and RF) could be slightly distinguished from each other along the first axis (Fig. 1a). Individuals from the second site (SB) appeared intermixed between these two clusters, without clear genetic differentiation, although some individuals separated along the second axis. When surveying if individuals clustered by age class, no obvious groups were observed, suggesting no relationship between age class and genetic relatedness (Fig. 1b). No further differences were observed among age classes in average or distribution of minor allele frequencies (Fig. S2A and B). Pairwise F_{ST} statistics between groups were low and quantitatively homogeneous (between 0.0103 and 0.0114; Table S3). Pairwise F_{ST} among age classes were also low (between 0.0008 and 0.0145; Table S4), with the highest values observed between 6- and 12-years-old individuals within SB (Table S1).

The pairwise genomic relatedness matrix (G; Fig. S3) revealed only small subsets of closely related individuals within groups. The spectral decomposition of this matrix revealed that less than 5% of the genetic variation was

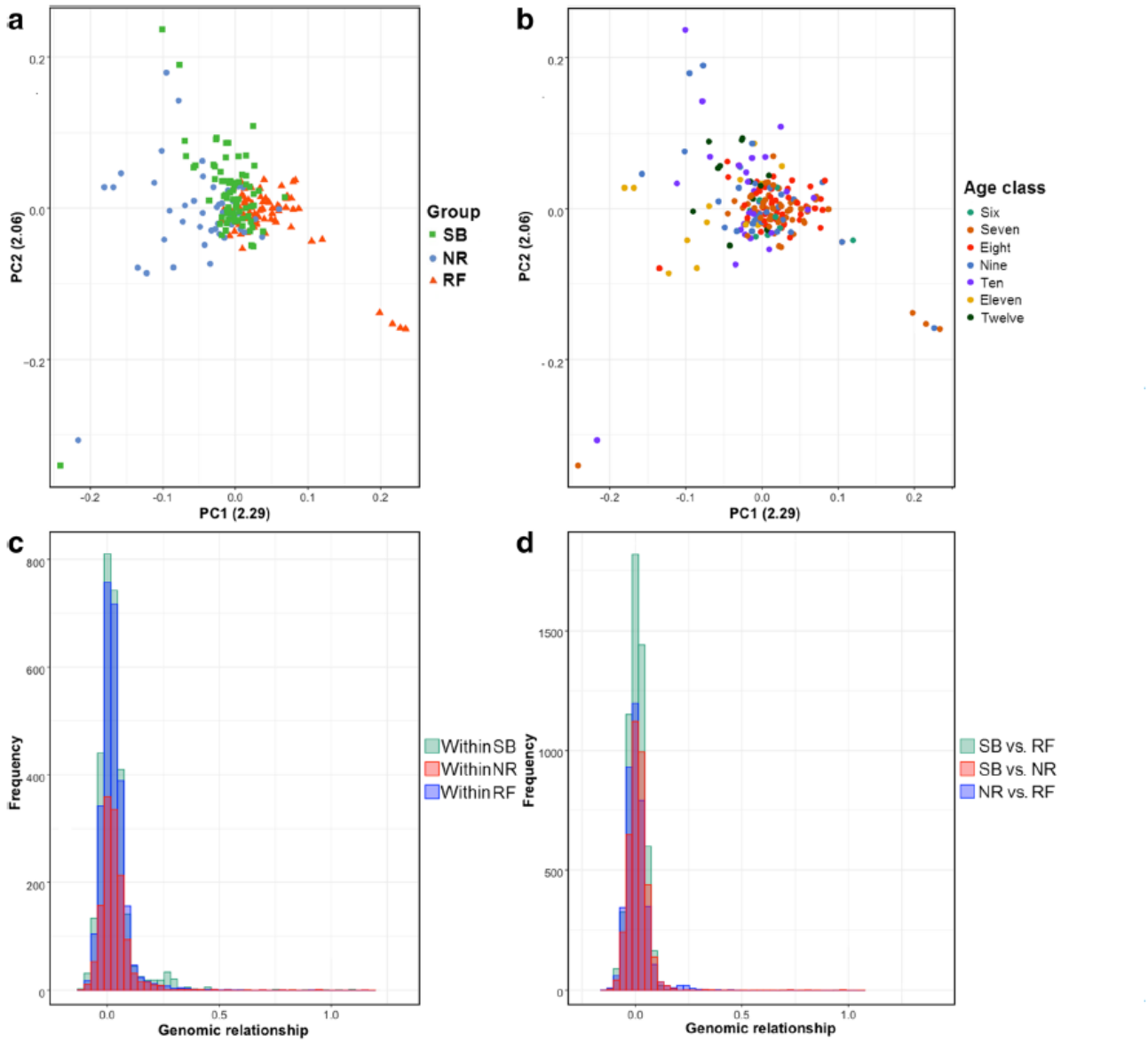

Fig. 1 **a**, **b** First two principal components of a PCA on 2286 SNP genotypes for sacred fir (*Abies religiosa*) individuals sampled in two trial sites in central Mexico (San Bartolo and Rincón de Guadalupe). Colors denote origin groups (**a**) and age classes (**b**; by year) (see color palates on the right). Numbers in brackets along axes indicate % of explained genotypic variance. **c**, **d** Distribution of pairwise relationships between individuals of a same group (**c**; excluding diagonal elements), and of different groups (**d**). SB, natural regeneration at San Bartolo; NR, natural regeneration at Rincón de Guadalupe; RF, introduced saplings at Rincón de Guadalupe

captured by the first eigenvector (4.62), indicating that population structure was low and that G was adequately capturing between-individuals relationships (Beaulieu et al. 2014b). Among the 201 genotyped trees, approximately 15.2, 10.6, and 9.3% of the individuals were half-cousins at NR, SB, and RF, respectively; 3.5, 2.2, and 2.8% of trees within the same groups were first cousins; 1.9, 1.6, and 3.1% were half-sibs, and approximately 0.6, 0.4, and 0.5% were full-sibs, respectively (Fig. 1c). The paired-relatedness values between individuals of different groups revealed virtually no full-sibs and only between 0.32% (SB vs. NR) and 1.14% (NR vs. RF) of half-sibs (Fig. 1d); approximately 6.5, 7.0, and 5.1% of the individuals were half-cousins between the SB vs. RF, SB vs. NR, and NR vs. RF groups, and 0.2, 1.0, and 1.5% were first-cousins, respectively. Mean pairwise linkage disequilibrium within-groups, according to the LDN_e method, was low ($r^2 < 0.2$), averaging 0.029, 0.021, and 0.017 for NR, RF, and SB, respectively (Fig. S4). Although there were more pairs of markers exhibiting above-average LD values for the group of reforested individuals (RF) than for the naturally regenerated groups (NR and SB). The estimated effective population sizes (N_e) based on these LD data were respectively 43, 56, and 122, for NR, RF, and SB, which is similar to sample sizes used (51 for NR, 77 for RF, and 73 for SB).

Predictive ability and genomic heritability

When considering the combined model for the whole sample, the two single-trait genomic prediction (GP) methods used (BRR and RKHS) had a similar predictive ability for each individual trait (Fig. 2); although BRR had the lowest standard errors (Table S5) and was judged more reliable than RKHS. Predictive abilities were low to moderate, being higher for growth (between 0.25 and 0.46) than for physiological traits (between 0.08 and 0.35) ($p < 0.05$; Table 3 and Table S6). Among traits, narrow sense heritability (h^2) ranged between 0.12 for RGR and 0.29 for G_{2016}. Growth traits had significant higher h^2 values (from 0.18 to 0.29; Table 3) than physiological traits (from 0.12 to 0.24; Table 3). Predictive ability and heritability estimates for groups of individuals were significantly higher for the naturally regenerated individuals at both sites ($r_y = -0.1$–0.43; and $h^2 = 0.18$–0.35; Table 3 and Table S7), especially for SB plants, than those for the reforested individuals (r_y values from -0.05 to 0.36; and $h^2 = 0.16$–0.29; Table 3 and Table S7).

Cross-validation among groups (training group (G_{TRN}) and test group (G_{TST})) with four types of models and training/testing dataset combinations showed that model accuracy improved as TRN/TST ratio increased for the four model types and training/testing dataset combinations (Fig. 3; Table S8). Results revealed that the highest predictive ability was achieved when models were calibrated with naturally regenerated individuals and used to predict phenotypes from introduced plants in the same site (Fig. 3; NR → RF; r_y values between 0.13 for RGR and 0.38 for G_{2016}). The lowest predictive ability was obtained when phenotypes of reforested individuals were predicted from models built for plants from a different location (SB → RF; r_y values from 0.08 for RWC to 0.23 for G_{2016}). Similarly, predicting phenotypes of naturally regenerated individuals from one site using models built for a different site (NR → SB and SB → NR) also provided relatively low values, with predictive ability being better for the first combination than for the second one (Fig. 3). This implies that the relationships captured by the models only allowed them to perform relatively well for the same environment and, thus, that our design did not allow capturing genotype-by-environment interactions, which reduced predictive ability between sites.

Interestingly, r_y increased for various traits (G_{2016}, AGB, CR, RWC, SD and Ψ; Tables S2 and S8) when including between 30 and 40% of individuals from the target group (the one to be predicted) in the training set, mainly for the NR → RF and NR → SB models (Fig. 3). Thus, allowing for some transferability. With such figures (i.e., 74 individual of TRN and 54 individuals of TST; Table S2), the best model (NR → RF) provided the highest predictive abilities for growth during the previous growing season (r_{G2016}, = 0.36), water potential ($r_\Psi = 0.27$), stem diameter ($r_{SD} = 0.26$), and aboveground biomass ($r_{AGB} = 0.24$) (Tables S2 and S8), which were also the traits that increased more constantly. These traits were retained as top traits for further analyses.

Most predictive SNP datasets for top traits

Models and combinations of individuals maximizing predictive ability (NR → RF) were used for optimizing predictive SNP datasets for the four top traits above by forming subsets of SNPs with a progressive increase of marker effects. Predictive ability (r_y) increased with the number of SNPs until reaching a plateau at approximately 1200 markers (Fig. 4). Optimized datasets comprised between 700 and 900 SNPs per trait and produced similar r_y values than the full dataset of 2286 SNPs. Interestingly, the optimized dataset for predicting Ψ (a physiological trait) included less genetic markers (700) than those for growth characters (800, 800,

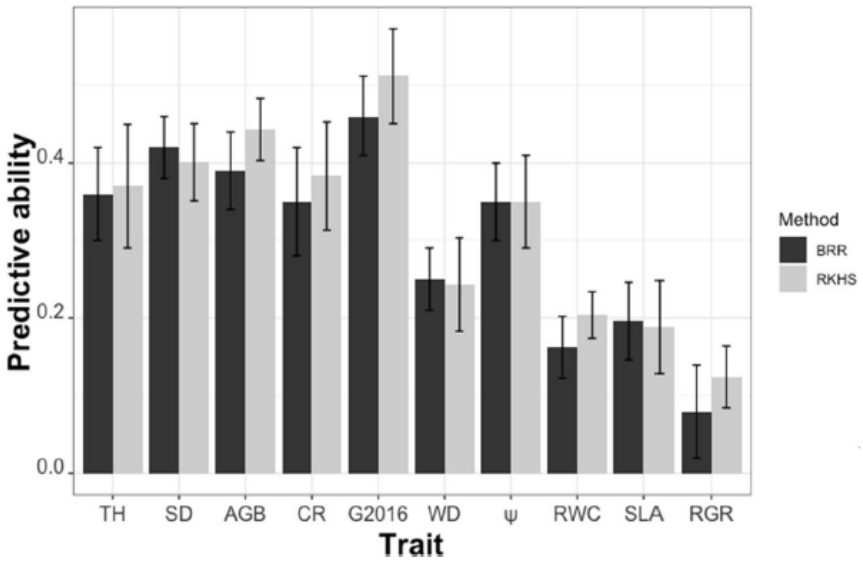

Fig. 2 Predictive ability (r_y) determined from 20-fold cross-validation with two single-trait genomic selection models (RKHS and BRR) performed for the whole sample of sacred fir (*Abies religiosa*) individuals. Brackets within bars denote standard errors. Abbreviations: TH, total height; SD, stem diameter; AGB, aboveground biomass; CR, crown ratio; G_{2016}, seasonal growth at 2016; WD, wood density; Ψ, water potential; RWH, needle relative water content; SLA, specific foliar area; RGR, relative growth rate. See Table S5 for a complete description of estimates per trait for each method

Table 3 Parameters of genomic prediction (GP) models after 20-fold cross-validation for ten phenotypic traits evaluated for three groups (SB, NR, and RF) and the whole sample of *A. religiosa* individuals from two trial sites at central Mexico: natural regeneration at San Bartolo (SB, 73 individuals), Rincon de Guadalupe (NR, 51 individuals), and introduced saplings (RF, 77 individuals). Estimates of genomic heritability (h^2) and predictive ability (r_y) are followed by their respective standard errors (*SE*)

Phenotypic trait	Whole sample				San Bartolo				Rincon de Guadalupe							
					Natural regeneration (SB)				Natural regeneration (NR)				Reforestation (RF)			
	h^2	SE	r_y	SE	h^2	SE	r_y	SE	h^2	SE	r_y	SE	h^2	SE	r_y	SE
Growth traits																
Total height (TH)	0.23	0.03	0.36	0.06	0.30	0.04	0.42	0.05	0.26	0.05	0.33	0.07	0.21	0.04	0.22	0.08
Stem diameter (SD)	0.26	0.03	0.42	0.04	0.32	0.06	0.33	0.07	0.27	0.04	0.36	0.08	0.25	0.06	0.27	0.06
Aboveground biomass (AGB)	0.25	0.04	0.39	0.05	0.31	0.05	0.32	0.06	0.25	0.05	0.33	0.09	0.23	0.05	0.25	0.07
Crown ratio (CR)	0.24	0.04	0.35	0.07	0.35	0.05	0.43	0.06	0.24	0.04	0.25	0.10	0.23	0.05	0.20	0.06
2016's growth (G_{2016})	0.29	0.04	0.46	0.05	0.31	0.06	0.37	0.07	0.31	0.06	0.42	0.09	0.29	0.06	0.36	0.05
Wood density (WD)	0.18	0.05	0.25	0.04	0.29	0.05	0.24	0.06	0.22	0.04	0.23	0.09	0.18	0.04	0.10	0.06
Physiological traits																
Water potential (Ψ)	0.24	0.03	0.35	0.05	0.30	0.05	0.33	0.07	0.26	0.06	0.35	0.08	0.24	0.04	0.34	0.06
Relative water content (RWC)	0.16	0.04	0.16	0.04	0.26	0.06	0.19	0.08	0.23	0.05	0.15	0.07	0.22	0.05	0.08	0.06
Specific leaf area (SLA)	0.17	0.03	0.20	0.05	0.27	0.05	0.22	0.08	0.21	0.05	0.18	0.08	0.21	0.06	0.15	0.07
Relative growth rate (RGR)	0.12	0.02	0.08	0.06	0.21	0.03	−0.1	0.05	0.18	0.05	−0.003	0.06	0.16	0.03	−0.05	0.08

and 900 for SD, AGB, and G_{2016}, respectively). Only 98 out 2286 SNPs were shared among optimized models (Fig. 5). More SNPs (135) were shared between optimized models for growth traits (SD, AGB, and G_{2016}) than between models for the physiological top trait (Ψ) and any of the top growth traits. Optimized models for G_{2016} had the highest number of unique SNPs (223), while SD and AGB had the lowest numbers (157 and 156, respectively).

Discussion

In this pilot study, we have shown that GP models can be used to predict some phenotypic traits in young individuals from small and secluded natural tree populations. By testing different training and validation datasets in across-groups models, we obtained acceptable predictive abilities for growth during the previous growing season (r_{G2016} = 0.36), water potential (r_Ψ = 0.27), stem diameter (r_{SD} = 0.26), and aboveground biomass (r_{AGB} = 0.24). The best models were those trained with natural individuals and used to predict the phenotype of introduced plants. In some cases, predictive ability could be achieved with over 700 SNPs, particularly for Ψ. These results offer new insights to continue exploring GP in non-model tree species and open a possible expansion to small and secluded natural populations. If confirmed, GP models could be eventually used for guiding management and conservation efforts, particularly for species with restricted distributions and/or key populations in the "rear-edge" of species' ranges.

Feasibility of GP modeling in small secluded natural populations

Estimated genomic heritabilities (h^2) and predictive abilities (r_y) were similar to those reported for breeding populations of boreal conifers (Resende et al. 2012b; Li et al. 2019; Calleja-Rodriguez et al. 2020), which are naturally distributed in much larger and interconnected populations (and with probably higher effective population sizes) than their tropical counterparts, including sacred fir. In spite of the low sample size and genome coverage used herein, our pilot-scale results are preliminary but encouraging and call for further research efforts for a better development of site-specific GP models in natural stands.

Inferences at small local scales, such as performed herein, may be aided by weak genetic structure and increased genetic relatedness that has been building up over various generations of inbreeding (Sedlacek et al. 2016). This would be increasing both LD between causal and genotyped markers, and the proportion of shared alleles among individuals; they both would allow capturing large fractions of the underlying genetic variation without the need of a structured pedigree. This is somehow contrary to what has been observed in open pollinated populations of widespread conifers like boreal spruces (Lenz et al. 2017), for which rampant gene flow may be diminishing LD and thus prediction ability. Incoming gene flow in sacred fir is actually expected to be low, as evidenced by its strong population differentiation (Cruz-

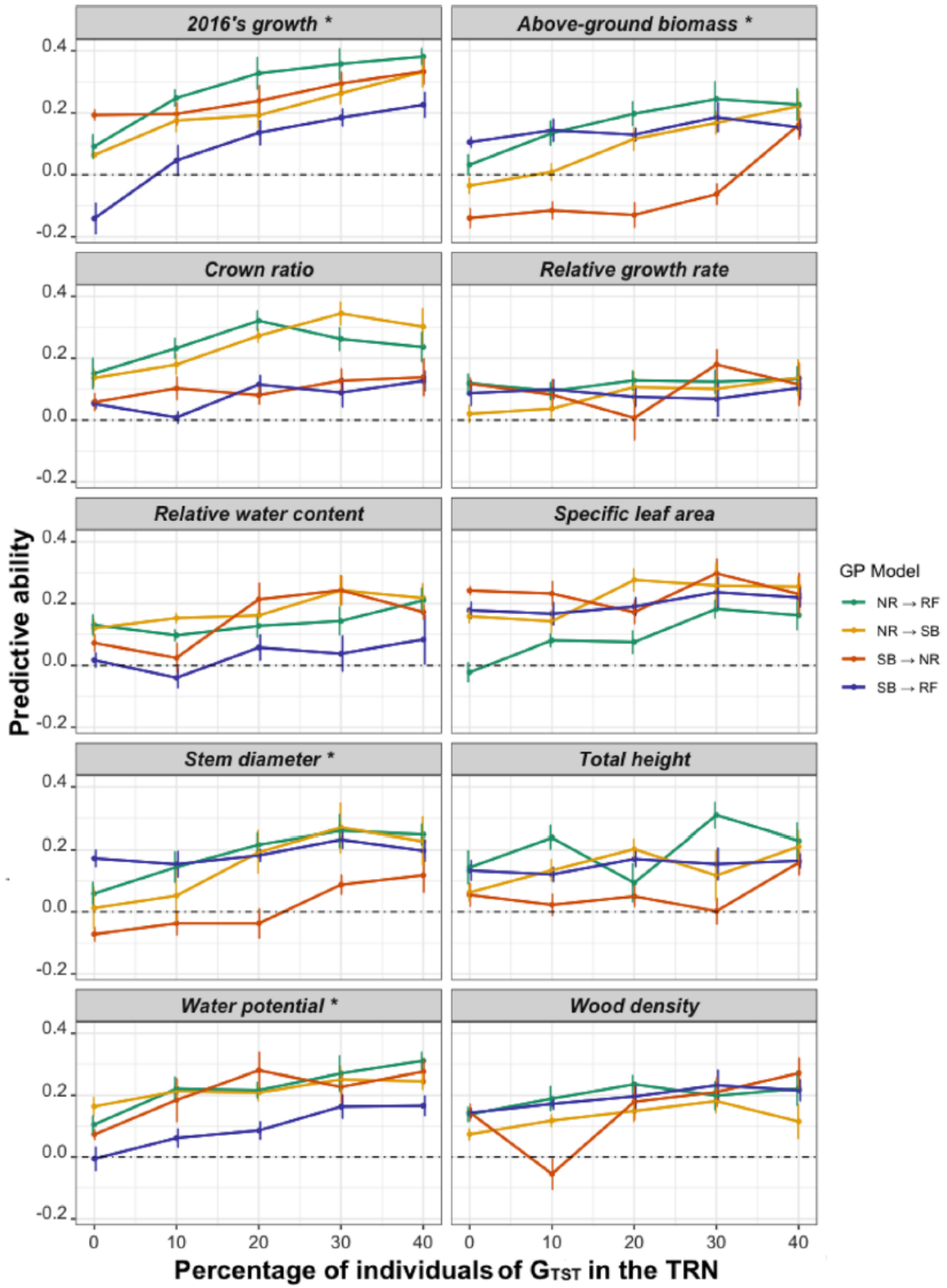

Fig. 3 Predictive ability (r_y), after 20-fold CV, for different predictive models "across-groups" (colored lines) with increasing percentages of individuals of the target group (training group; G_{TST}) included in the training set (TRN). Only the BRR method was performed. Retained "top traits" are marked with asterisks. The composition of each set can be seen in Table S2; a complete description of model performance for each trait is available in Table S8. Dashed lines indicate zero values for r_y

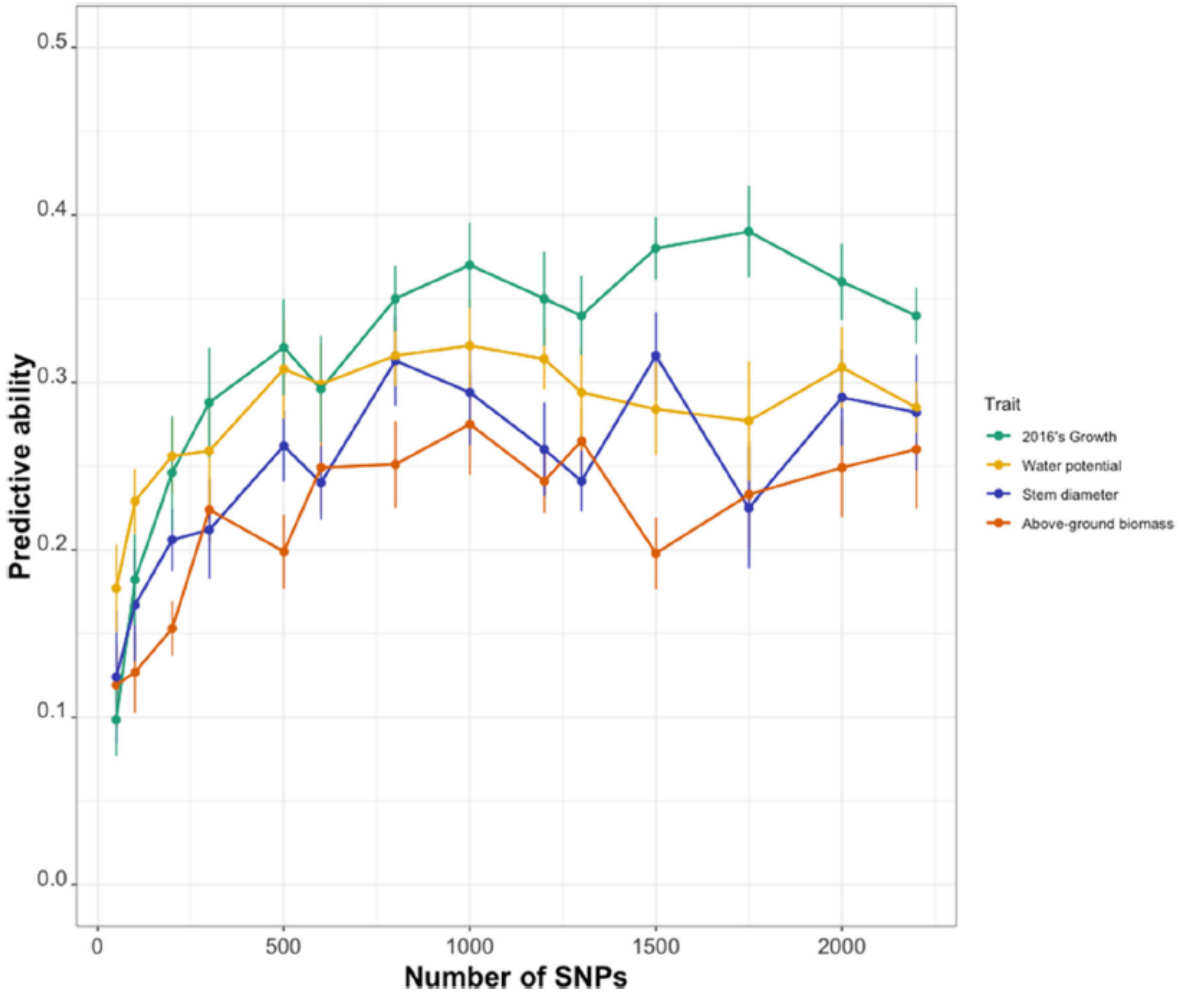

Fig. 4 Effect of the number of SNPs on predictive ability (r_y) for four "top traits" (colored lines) in sacred fir (*Abies religiosa*). Models built for naturally reforested individuals were used to predict performance of reforested plants (i.e., NR → RF model with BRR method). r_y was assessed with random 20-fold cross-validations for selecting markers. Subsets include SNPs with decreasing marker effect. See text for details

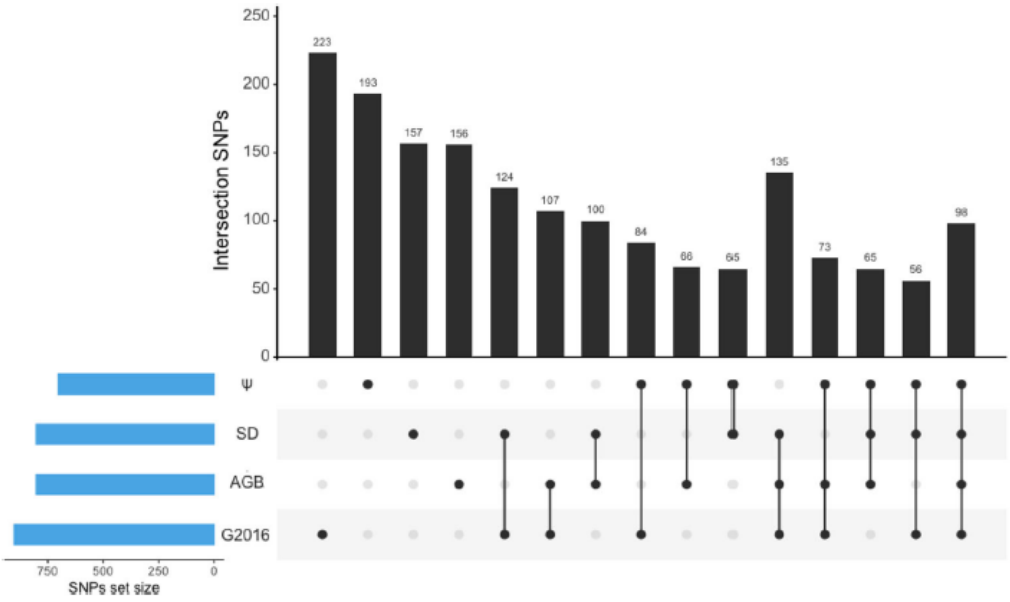

Fig. 5 Most predictive SNP datasets for four "top traits" (lower panel) in sacred fir (*Abies religiosa*). Upset chart shows the number of SNPs per trait (individual dots below chart) and overlap between SNP subsets between traits (connected dots below chart). SNPs were sorted up according to predictive ability. Size for the "optimum" SNP set for predicting each "top trait" is indicated on the horizontal bar plot on the lower left corner

Nicolás et al. 2020), which is sometimes even observed at small geographic scales (Méndez-González et al. 2017).

GP is based on the assumption that genomic h^2 values reflect true genetic relationships among individuals and, as such, correspond to the proportion of phenotypic variance that can be explained by a regression against molecular marker variation; this, at its turn, is conditioned by N_e and the magnitude of LD (Chen et al. 2018; Lebedev et al. 2020; Thistlethwaite et al. 2020). Indeed, previous simulations (Grattapaglia and Resende 2011) have indicated that both low N_e and increased LD may result in fair predictive accuracies even when h^2 and the number of individuals in the training set are modest. Our N_e estimates, together with the particular demographic history and secluded nature of the stands surveyed, indicate small historical effective population sizes, and populations governed by strong genetic drift, which is typical of high mountain stands (e.g., Robledo-Arnuncio et al. 2004; Cobo-Simón et al. 2020). All these factors should be increasing non-random associations (LD) between markers and QTLs (Rogers 2014; Thistlethwaite et al. 2019), although our pilot study likely captured only a small part of these associations because of low marker density. Using larger datasets, both in terms of sample size and marker density, together with realized relationship models (BLUPs) should allow for a better estimation of quantitative genetic parameters, such as performed for populations with shallow pedigree structure in white spruce (Gamal El-Dien et al. 2016).

As previously observed in other forest trees, each marker only explained but a small fraction of the heritable phenotypic variance in our pilot study, which is consistent with a polygenic architecture for the studied traits (Mitchell-Olds et al. 2007; Eckert et al. 2013; Barghi et al. 2020). Narrow sense genomic heritability and predictive ability were generally higher for growth (particularity for G_{2016}, SD, and AGB for three groups and combined model) than for physiological traits ($p < 0.05$; Table S6), thus suggesting fewer genetic variants underlying phenotypic variation for the former than for later traits (Bouvet et al. 2020). Indeed, a higher proportion of genes and biologically relevant pathways should be influencing physiological traits (Plomion et al. 2016; Moran et al. 2017), which may fit more adequately an infinitesimal genomic model (Müller et al. 2017) than the one used herein. The lower h^2 and r_y estimates for physiological traits may also be reflecting their predominantly plastic nature, driven by significant genotype × environment (G × E) and epistatic components (Gage et al. 2017).

The characteristic seasonal nature of rain and drought periods in central Mexico, together with their recurrent interannual variations (Karmalkar et al. 2011), must be further noted, which renders extremely challenging obtaining a unique, precise and adequate measure of water use efficiency in plants of this region. Such a challenge further increases the difficulty for establishing selection criteria for genotypes with enhanced hydric potential and which could be better pre-adapted to water limitations. Addressing these difficulties should be a priority for future studies.

Enhancing GP predictive ability for natural populations

Cross-validation indicated that GP model performance was poor when exclusively using individuals from one group (G_{TST}) to predict the phenotypes of another group (G_{TRN}). Indeed, models only performed relatively well when the TRN included between 30 and 40% of individuals from the target (i.e., testing) group (Fig. 3). This might be a factor to consider when implementing GP in management programs involving introduction of foreign material and a top priority for future studies. Indeed, to our knowledge, this has not been addressed in previous surveys, which only indicated that predictive ability increases with TRN/TST ratio (e.g., Tan et al. 2017; Calleja-Rodriguez et al. 2020) and with larger numbers of trees per family (Müller et al. 2017; Chen et al. 2018). Understanding how this may be applied to model building in natural secluded populations is still a pending task. We hypothesize that effectively sampling the whole relationship spectrum within the population of interest, including the contemporary genetic relatedness and population structure between the TRN and TST sets, should be key for optimizing predictability in natural stands. This represents a significant challenge, as GP model success in natural stands would indirectly rely in the amount of regeneration and establishment success, which is often modest for slow-growing species inhabiting closed tropical forest stands. Indeed, this was the main limiting factor for sample size in our pilot study.

In any case, the model built from naturally regenerated individuals and used to predict phenotypes of introduced plants in the same site (NR → RF) was the one that best captured effects for traits like G_{2016}, TH, and SD (Fig. 3), indicating that transferring GP models between environments might not be possible, even when increasing admixture in the training set. Indeed, as in previous studies, phenotypic prediction diminished when forecasting the performance of introduced plants in one location from models built for individuals from another site (i.e., SB → RF model). Interestingly, this was observed for relatively closely located stands (7.4 km from each other), which must be particularly compelling for reforestation or assisted migration programs into locations well outside the natural species' range, or into sites where natural populations no longer occur (Carbajal-Navarro et al. 2019). In such circumstances, GP may have limited predictive ability and other strategies should be employed; for instance, establishing a trial test first and then build GP models from phenotypic variation on that test. Another solution could be building multi-population GP models (e.g., Beaulieu et al. 2020) that include a much larger number of individuals

originating from highly differentiated environments and elevations. This would allow accounting for genotype-phenotype links among stands, and taking G × E interactions into account, which may improve model accuracy, as shown with cross-site models in various forest trees (e.g., Resende et al. 2012a; Lenz et al. 2017). Integrating such interactions appears as an additional challenge to address when applying GP in natural populations.

The effect of SNP number on genomic prediction of traits

Besides the composition and size of the training sets, the number of SNPs is an important factor that influences genomic prediction success (Tan et al. 2017; Calleja-Rodríguez et al. 2020; Thistlethwaite et al. 2020). Higher marker density typically increases marker-QTL LD, and therefore r_y estimates (Grattapaglia and Resende 2011; Desta and Ortiz 2014). Yet, our results were consistent with early pilot studies exploring GP feasibility in conifers (Resende et al. 2012a; Zapata-Valenzuela et al. 2012; Zapata-Valenzuela et al. 2013), indicating that small SNP subsets (≈ 1 K) chosen according to size effect (β) may also recover enough genetic relationships that allow capturing some predictive ability. In our case, this is the likely result of working with small secluded natural populations, for which LD might be higher because of historical consanguinity and genetic drift. Another possibility is that high-effect markers were captured, which provided sufficient information to trace back genotype-phenotype relationships, although additional surveys would be needed for prove this. In any case, our results suggest that low-density SNP chips could be considered for reducing genotyping costs without risking diminished prediction ability, as routinely employed in domestic animals (Habier et al. 2010; Van Eenennaam et al. 2014). Exploring such an avenue would facilitate practical applications of GP in natural forest tree populations, like in situ management and conservation, and assisted migration.

Perspectives

Selection of individuals with faster growth rates and higher yield is at the core of modern GP-assisted tree breeding programs (Grattapaglia et al. 2018; Cortés et al. 2020). However, fast growth often implies an ecological trade-off for reduced water stress tolerance (Resco de Dios et al. 2018; Redmond et al. 2019). Adopting GP for traits related to water use efficiency is one of the main challenges for forecasting adaptation to climate change, particularly because the genomic bases of these traits remain poorly understood (Bouvet et al. 2020; Lebedev et al. 2020). Forest ecosystems, including plantations, are currently subjected to higher drought pressures than in the past, because of increasing temperatures and changes in precipitation regimes (Choat et al. 2012; Schuster and Oberhuber 2013; Wachowiak et al. 2018). Therefore, preselecting genotypes for mitigating hydraulic failures and improve water use efficiency (net carbon fixed per unit water) would need targeting genes or pathways involved in fast-acting physiological mechanisms that allow for more stable Ψ's over individuals' life-spans. This seems particularly difficult for natural populations, where estimating such parameters for large samples and under uncontrolled conditions is even more challenging.

Following the results of this pilot study, evaluating GP performance in juvenile individuals from small secluded populations should also be a priority, especially for populations in the "rear edge" of species' distributions. These "rear edge" stands are often seen as reservoirs of adaptive variants to future climate conditions (Hampe and Petit 2005). Thus, integrating GP models to account for such variants could ease our understanding of population resilience, and help planning future management efforts. However, such GP models might need adjustment for spatial variation that corrects for environmental structure, heterogeneity, and G × E (Lenz et al. 2020b). "Speed breeding" and machine learning (Libbrecht and Noble 2015) techniques may help integrating all these factors, while remote sensing (Dungey et al. 2018) and "scale-free" climate algorithms (ClimateBC, Wang et al. 2012) will allow for better capturing environmental heterogeneity (Ratcliffe et al. 2019).

Further pending questions include how well GP models will perform in natural populations over time, how often these models must be updated, and if they would benefit from newly gathered phenotypic data that account for allele frequency changes across generations (Grattapaglia 2014). Answering these and other questions would validate the effectiveness of GP as a tool for managing natural forests and increasing tree performance and survival. Such validation will be the first step for large-scale programs merging different strategies; for example, GP-guided relocations relying on machine learning approaches for better integrating genotype-phenotype data (González-Camacho et al. 2018). Such an approach should be more efficient than pre-selecting individuals or source populations for reforestation based solely on allele frequencies at candidate loci.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11295-020-01489-1.

Acknowledgments We express our gratitude to the "ejidatarios" from Rincón de Guadalupe and San Bartolo and Andrés Juarez, Eusebio Roldán, and Lucia Madrid at the "Consejo Civil Mexicano para la Sivilcultura Sotenible" for granting access to the forest trials and sharing their knowledge. We also thank Gustavo Giles, Veronica Reyes, Jorge Cruz, Alfredo Villarruel, Karen Carrasco, and Armando Sunny for fieldwork assistance, and Tania Garrido, Nancy Gálvez, Laura Giraldo, and Azalea Guerra for valuable help in laboratory analyses. We are grateful to Ernesto Campos, Felipe López-Hernández, Leopoldo Vázquez, and

Gustavo Giles for assistance in bioinformatics analyses. We additionally thank the Department of Environmental and Soil Sciences of the Institute of Geology (UNAM) for edaphic analyses and the Functional Ecology (IE-UNAM) and Tissue Culture (Jardín Botánico-UNAM) laboratories at UNAM, and the Department of Forest Sciences at Universidad Nacional de Colombia for logistic support. Bioinformatic analyses were performed on a computing cluster at the Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), which is supported by the system administrator and the subcoordinación de soporte informático at CONABIO. This paper is part of the doctoral research conducted by SA, who thanks the "Posgrado en Ciencias Biológicas at the Universidad Nacional Autónoma de México" and acknowledges a scholarship from the "Consejo Nacional de Ciencia y Tecnología (CONACyT; grant no. 480152)".

Data archiving statement Data will be available in the FigShare data repository (https://figshare.com/) upon manuscript acceptance. Access links will be provided when available.

Author's contributions SA, AM-Y, and JPJ-C designed the study. SA and AM-Y performed fieldwork. SA carried out molecular and phenotypic analyses. SA and AJC performed statistical analyses and interpreted results. SA and JPJ-C drafted the manuscript. All authors reviewed and approved the final manuscript.

Funding This work was financially supported by grants from CONACyT (CB-2016-284457 and 278987) and both the "Dirección General de Asuntos del Personal Académico" (PAPIIT project: IN208416) and the Institute of Ecology (presupuesto operativo) at UNAM to JPJ-C.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

Submission declaration and verification Authors declare that the work here described has not been previously published and is not under consideration for publication elsewhere. In addition, its publication is approved by all authors and the responsible authorities where the field work was carried out (Consejo Civil Mexicano para la Silvicultura Sostenible), and that, if accepted, it will not be published elsewhere in the same form, in English or any other language, including electronically.

References

Abramoff MD, Magalhães PJ, RamSJ (2004) Image processing with ImageJ. Biophoton Int 11(7):36–42

Aitken SN, Bemmels JB (2015) Time to get moving: assisted gene flow of forest trees. Evol Appl 9(1):271–290. https://doi.org/10.1111/eva.12293

Aitken SN, Whitlock MC (2013) Assisted gene flow to facilitate local adaptation to climate change. Annu Rev Ecol Evol Syst 44:367–388. https://doi.org/10.1146/annurev-ecolsys-110512-135747

Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. Evol Appl 1:95–111. https://doi.org/10.1111/j.1752-4571.2007.00013.x

Amadeu R, Cellon C, Lara L, Resende M, Oliveira I, Ferrao L, Munoz P, Garcia A (2016) Package 'AGHmatrix' R package to construct relationship matrices for autotetraploid and diploid species: a blueberry example. The Plant Genome 9. https://doi.org/10.3835/plantgenome2016.01.0009

Barghi N, Hermisson J, Schlötterer C (2020) Polygenic adaptation: a unifying framework to understand positive selection. Nat Rev Genet 21:769–781. https://doi.org/10.1038/s41576-020-0250-z

Barton NH, Etheridge AM, Véber A (2017) The infinitesimal model: definition, derivation, and implications. Theor Popul Biol 118:50–73. https://doi.org/10.1016/j.tpb.2017.06.001

Beaulieu J, Doerksen T, Clément S, Mackay J, Bousquet J (2014a) Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. Heredity (Edinb) 113:343–352. https://doi.org/10.1038/hdy.2014.36

Beaulieu J, Doerksen TK, Mackay J, Rainville A, Bousquet J (2014b) Genomic selection accuracies within and between environments and small breeding groups in white spruce. BMC Genomics 15(1048):1–16

Beaulieu J, Nadeau S, Ding C, Celedon JM, Azaiez A, Ritland C, Laverdière JP, Deslauriers M, Adams G, Fullarton M, Bohlmann J, Lenz P, Bousquet J (2020) Genomic selection for resistance to spruce budworm in white spruce and relationships with growth and wood quality traits. Evol Appl 13:2704–2722. https://doi.org/10.1111/eva.13076

Bezemer N, Krauss SL, Roberts DG, Hopper SD (2019) Conservation of old individual trees and small populations is integral to maintain species' genetic diversity of a historically fragmented woody perennial. Mol Ecol 28(14):3339–3357. https://doi.org/10.1111/mec.15164

Bouvet JM, Makouanzi-Ekomono CG, Brendel O, Laclau JP, Bouillet JP, Epron D (2020) Selecting for water use efficiency, wood chemical traits and biomass with genomic selection in a Eucalyptus breeding program. For Ecol Manag 465:118092. https://doi.org/10.1016/j.foreco.2020.118092

Bradbury PJ, Zhang DE, Kroon TM, Casstevens Y, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23(19):2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Brian A, Peterson BG, Carl P, Boudt K, Bennett R, Ulrich J et al (2019) PerformanceAnalytics: econometric tools for performance and risk analysis. R package version 0.9 9 **(2)**

Caldwell KS, Russell J, Langridge P, Powell W (2006) Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. Genetics 172(1):557–567. https://doi.org/10.1534/genetics.104.038489

Calleja-Rodriguez A, Pan J, Funda T, Chen Z, Baison J, Isik F, Abrahamsson S, Wu HX (2020) Evaluation of the efficiency of genomic versus pedigree predictions for growth and wood quality traits in Scots pine. BMC Genomics 21:1–17. https://doi.org/10.1186/s12864-020-07188-4

Campoy JA, Lerigoleur-Balsemin E, Christmann H, Beauvieux R, Girollet N, Quero-García J, Dirlewanger E, Barreneche T (2016) Genetic diversity, linkage disequilibrium, population structure and construction of a central collection of local races and improved cultivars of Prunus avium L. BMC Plant Biol 16:49. https://doi.org/10.1186/s12870-016-0712-9

Cappa EP, Stoehr MU, Xie CY, Yanchuk AD (2016) Identification and joint modeling of competition effects and environmental heterogeneity in three Douglas-fir (*Pseudotsuga menziesii* var. menziesii) trials. Tree Genet Genomes 12:102. https://doi.org/10.1007/s11295-016-1061-4

Carbajal-Navarro A, Navarro-Miranda E, Blanco-García A, Cruzado-Vargas AL, Gómez-Pineda E, Zamora-Sánchez C et al (2019) Ecological restoration of *Abies religiosa* forests using nurse plants and assisted migration in the Monarch Butterfly Biosphere Reserve, Mexico. Front Ecol. Evol 7. https://doi.org/10.3389/fevo.2019.00421

Chen ZQ, Baison J, Pan J, Karlsson B, Andersson B, Westin J, García-Gil MR, Wu HX (2018) Accuracy of genomic selection for growth and wood quality traits in two control-pollinated progeny trials using

exome capture as the genotyping platform in Norway spruce. BMC Genomics 19(946):946. https://doi.org/10.1186/s12864-018-5256-y

Choat B, Jansen S, Brodribb TJ, Cochard H, Delzon S, Bhaskar R, Bucci SJ, Feild TS, Gleason SM, Hacke UG, Jacobsen AL, Lens F, Maherali H, Martínez-Vilalta J, Mayr S, Mencuccini M, Mitchell PJ, Nardini A, Pittermann J, Pratt RB, Sperry JS, Westoby M, Wright IJ, Zanne AE (2012) Global convergence in the vulnerability of forests to drought. Nature 491:752–755. https://doi.org/10.1038/nature11688

Chojnacky DC, Heath LS, Jenkins JC (2014) Updated generalized biomass equations for North American tree species. Forestry 87:129–151. https://doi.org/10.1093/forestry/cpt053

Chung MY, Epperson BK, Chung MG (2003) Genetic structure of age classes in *Camellia japonica* (Theaceae). Evolution 57(1):62–73. https://doi.org/10.1554/0014-3820(2003)057[0062:GSOACI]2.0.CO;2

Cobo-Simón I, Méndez-Cea B, Jump AS, Seco J, Gallego FJ, Linares JC (2020) Understanding genetic diversity of relict forests. Linking long-term isolation legacies and current habitat fragmentation in Abies pinsapo Boiss. For Ecol Manag 461:117947. https://doi.org/10.1016/j.foreco.2020.117947

Cochard H, Cruiziat P, Tyree MT (1992) Use of positive pressures to establish vulnerability curves: further support for the air-seeding hypothesis and implications for pressure-volume analysis. Plant Physiol 100(1):205–209. https://doi.org/10.1104/pp.100.1.205

Cortés AJ, Restrepo-Montoya M, Bedoya-Canas LE (2020) Modern strategies to assess and breed forest tree adaptation to changing climate. Front Plant Sci. https://doi.org/10.3389/fpls.2020.583323

Cruz-Nicolás J, Giles-Pérez G, González-linares E, Múgica-gallart J et al (2020) Contrasting stochastic and adaptive processes drive morphological and genetic differentiation in a subtropical fir (*Abies*, Pinaceae) species complex. Bot J Linn Soc 192(2):401–420. https://doi.org/10.1093/botlinnean/boz077

Cuevas J, Crossa J, Soberanis V, Pérez-Elizalde S, Pérez-Rodríguez P, de los Campos G, Montesinos-Lopez OA, Burgueño J (2016) Genomic prediction with genotype x environment interaction kernel regression models. The Plant Genome 9(3):1–20. https://doi.org/10.3835/plantgenome2016.03.0024

Dawson IK, Leakey R, Clement CR, Weber JC, Cornelius JP, Roshetko JM, Vinceti B, Kalinganire A, Tchoundjeu Z, Masters E, Jamnadass R (2014) The management of tree genetic resources and the livelihoods of rural communities in the tropics: non-timber forest products, smallholder agroforestry practices and tree commodity crops. For Ecol Manag 333:9–21. https://doi.org/10.1016/j.foreco.2014.01.021

de los Campos G, Sorensen D, Gianola D (2015) Genomic heritability: what is it? PLoS Genet 11(5):e1005048. https://doi.org/10.1371/journal.pgen.1005048

Desta ZA, Ortiz R (2014) Genomic selection: genome-wide prediction in plant improvement. Trends Plant Sci 19:592–601. https://doi.org/10.1016/j.tplants.2014.05.006

Dibdiakova J, Vadla K (2012) Basic density and moisture content of coniferous branches and wood in Northern Norway. In: McEvoy (ed) 2nd European Energy Conference, Maastricht, The Netherlands. EPJ Web of Conferences 33. https://doi.org/10.1051/epjconf/20123302005

Do C, Waples RS, Peel D, Macbeth GM, Tillett BJ, Ovenden JR (2014) Neestimator V2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. Mol Ecol Resour 14(1):209–214. https://doi.org/10.1111/1755-0998.12157

Dungey HS, Dash JP, Pont D, Clinton PW, Watt MS, Telfer EJ (2018) Phenotyping whole forests will help to track genetic performance. Trends Plant Sci 23:854–864. https://doi.org/10.1016/j.tplants.2018.08.005

Eaton DAR (2014) PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics 30:1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Eckert AJ, Wegrzyn JL, Liechty JD, Lee JM, Cumbie WP, Davis JM, Goldfarb B, Loopstra CA, Palle SR, Quesada T, Langley CH, Neale DB (2013) The evolutionary genetics of the genes underlying phenotypic associations for loblolly pine (*Pinus taeda*, Pinaceae). Genetics 195:1353–1372. https://doi.org/10.1534/genetics.113.157198

Finlay CMV, Bradley CR, Preston JS, Provan J (2017) Low genetic diversity and potential inbreeding in an isolated population of alder buckthorn (*Frangula alnus*) following a founder effect. Sci Rep 7:1–8. https://doi.org/10.1038/s41598-017-03166-1

Gage JL, Jarquin D, Romay C, Lorenz A, Buckler ES, Kaeppler S, Alkhalifah N, Bohn M, Campbell DA, Edwards J, Ertl D, Flint-Garcia S, Gardiner J, Good B, Hirsch CN, Holland J, Hooker DC, Knoll J, Kolkman J, Kruger G, Lauter N, Lawrence-Dill CJ, Lee E, Lynch J, Murray SC, Nelson R, Petzoldt J, Rocheford T, Schnable J, Schnable PS, Scully B, Smith M, Springer NM, Srinivasan S, Walton R, Weldekidan T, Wisser RJ, Xu W, Yu J, de Leon N (2017) The effect of artificial selection on phenotypic plasticity in maize. Nat Commun 8:1348. https://doi.org/10.1038/s41467-017-01450-2

Gamal El-Dien O, Ratcliffe B, Klápště J, Chen C, Porth I, El-Kassaby YA (2015) Prediction accuracies for growth and wood attributes of interior spruce in space using genotyping-by-sequencing. BMC Genomics 16:1–16. https://doi.org/10.1186/s12864-015-1597-y

Gamal El-Dien O, Ratcliffe B, Klápště J, Porth I, Chen C, El-Kassaby YA (2016) Implementation of the realized genomic relationship matrix to open-pollinated white spruce family testing for disentangling additive from nonadditive genetic effects. G3 (Bethesda) 6:743–753. doi: https://doi.org/10.1534/g3.115.025957

Garot E, Joët T, Combes MC, Lashermes P (2019) Genetic diversity and population divergences of an indigenous tree (*Coffea mauritiana*) in Reunion Island: role of climatic and geographical factors. Heredity (Edinb) 122:833–847. https://doi.org/10.1038/s41437-018-0168-9

Gentili R, Solari A, Diekmann M, Duprè C, Monti GS, Armiraglio S et al (2018) Genetic differentiation, local adaptation and phenotypic plasticity in fragmented populations of a rare forest herb. PeerJ 1–26. https://doi.org/10.7717/peerj.4929

George JP, Grabner M, Karanitsch-Ackerl S, Mayer K, Weißenbacher L, Schueler S (2017) Genetic variation, phenotypic stability, and repeatability of drought response in European larch throughout 50 years in a common garden experiment. Tree Physiol 37(1):33–46. https://doi.org/10.1093/treephys/tpw085

Gianola D, Fernando RL, Stella A (2006) Genomic-assisted prediction of genetic value with semiparametric procedures. Genetics. 173(3):1761–1776. https://doi.org/10.1534/genetics.105.049510

González-Camacho JM, Ornella L, Pérez-Rodríguez P, Gianola D, Dreisigacker S, Crossa J (2018) Applications of machine learning methods to genomic selection in breeding wheat for rust resistance. Plant Genome 11:1–15. https://doi.org/10.3835/plantgenome2017.11.0104

Goudet J, Jombart T (2015) Hierfstat: estimation and tests of hierarchical F-statistics. Mol Ecol Notes 5(1):184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x

Grattapaglia D (2014) Chapter 26: breeding forest trees by genomic selection: current progress and the way forward. In: Tubera et al (ed) Genomics of plant genetic resources: Volume 1. Managing, sequencing and mining genetic resources. Springerlink , pp 230–257. https://doi.org/10.1007/978-94-007-7572-5

Grattapaglia D, Resende MDV (2011) Genomic selection in forest tree breeding. Tree Genet Genomes 7:241–255. https://doi.org/10.1007/s11295-010-0328-4

Grattapaglia D, Silva-Junior OB, Resende RT, Cappa EP, Müller BSF, Tan B, Isik F, Ratcliffe B, el-Kassaby YA (2018) Quantitative

genetics and genomics converge to accelerate forest tree breeding. Front Plant Sci 871:1–10. https://doi.org/10.3389/fpls.2018.01693

Gutierrez AP, Matika O, Bean TP, Houston RD (2018) Genomic selection for growth traits in Pacific oyster (*Crassostrea gigas*): potential of low-density marker panels for breeding value prediction. Front Genet 9:1–9. https://doi.org/10.3389/fgene.2018.00391

Habier D, Tetens J, Seefried FR, Lichtner P, Thaller G (2010) The impact of genetic relationship information on genomic breeding values in German Holstein cattle. Genet Sel Evol 42:1–12. https://doi.org/10.1186/1297-9686-42-5

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. Ecol Lett 8(5):461–467. https://doi.org/10.1111/j.1461-0248.2005.00739.x

Hankin LE, Higuera PE, Davis KT, Dobrowski SZ (2018) Accuracy of node and bud-scar counts for aging two dominant conifers in western North America. For Ecol Manag 427:365–371. https://doi.org/10.1016/j.foreco.2018.06.001

Hoffmann WA, Poorter H (2002) Avoiding bias in calculations of relative growth rate. Ann Bot 90(1):37–42. https://doi.org/10.1093/aob/mcf140

Holliday JA, Zhou L, Bawa R, Zhang M, Oubida RW (2016) Evidence for extensive parallelism but divergent genomic architecture of adaptation along altitudinal and latitudinal gradients in *Populus trichocarpa*. New Phytol 209:1240–1251. https://doi.org/10.1111/nph.13643

Ingvarsson PK, Hvidsten TR, Street NR (2016) Towards integration of population and comparative genomics in forest trees. New Phytol 212:338–344. https://doi.org/10.1111/nph.14153

Isik F (2014) Genomic selection in forest tree breeding: the concept and an outlook to the future. New For 45(3):379–401. https://doi.org/10.1007/s11056-014-9422-z

Isik F, Bartholomé J, Farjat A, Chancerel E, Raffin A, Sanchez L, Plomion C, Bouffier L (2016) Plant science genomic selection in maritime pine. Plant Sci 242:108–119. https://doi.org/10.1016/j.plantsci.2015.08.006

Karmalkar A, Bradley RS, Diaz HF (2011) Climate change in Central America and Mexico: regional climate model validation and climate change projections. Clim Dyn 37:605–629

Klápště J, Dungey HS, Telfer EJ, Suontama M, Graham NJ, Li Y, McKinley R (2020) Marker selection in multivariate genomic prediction improves accuracy of low heritability traits. Front Genet 11:1–15. https://doi.org/10.3389/fgene.2020.499094

Koskela J, Vinceti B, Dvorak W, Bush D, Dawson IK, Loo J, Kjaer ED, Navarro C, Padolina C, Bordács S, Jamnadass R, Graudal L, Ramamonjisoa L (2014) Utilization and transfer of forest genetic resources: a global review. For Ecol Manag 333:22–34. https://doi.org/10.1016/j.foreco.2014.07.017

Larbi A, Morales F, Abadía J, Abadía A (2003) Effects of branch solid Fe sulphate implants on xylem sap composition in field-grown peach and pear: changes in Fe, organic anions and pH. J Plant Physiol 160:1473–1481. https://doi.org/10.1078/0176-1617-01010

Le HD, Smith C, Herbohn J, Harrison S (2012) More than just trees: assessing reforestation success in tropical developing countries. J Rural Stud 28:5–19. https://doi.org/10.1016/j.jrurstud.2011.07.006

Lebedev VG, Lebedeva TN, Chernodubov AI, Shestibratov KA (2020) Genomic selection for Forest tree improvement: methods, achievements and perspectives. Forests 11:1190. https://doi.org/10.3390/f11111190

Lenz PRN, Beaulieu J, Mansfield SD, Clément S, Desponts M, Bousquet J (2017) Factors affecting the accuracy of genomic selection for growth and wood quality traits in an advanced-breeding population of black spruce (*Picea mariana*). BMC Genomics 18:1–17. https://doi.org/10.1186/s12864-017-3715-5

Lenz PRN, Nadeau S, Mottet MJ, Perron M, Isabel N, Beaulieu J, Bousquet J (2020a) Multi-trait genomic selection for weevil resistance, growth, and wood quality in Norway spruce. Evol Appl 13:76–94. https://doi.org/10.1111/eva.12823

Lenz PRN, Nadeau S, Azaiez A, Gérardi S, Deslauriers M, Perron M, Isabel N, Beaulieu J, Bousquet J (2020b) Genomic prediction for hastening and improving efficiency of forward selection in conifer polycross mating designs: an example from white spruce. Heredity 124:562–578. https://doi.org/10.1038/s41437-019-0290-3

Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H, Manuscript A (2014) UpSet: visualization of intersecting sets Europe PMC funders group. IEEE Trans Vis Comput Graph 20(12):1983–1992. https://doi.org/10.1109/TVCG.2014.2346248

Li Y, Klápště J, Telfer E, Wilcox P, Graham N, Macdonald L, Dungey HS (2019) Genomic selection for non-key traits in radiata pine when the documented pedigree is corrected using DNA marker information. BMC Genomics 20:1026. https://doi.org/10.1186/s12864-019-6420-8

Libbrecht MW, Noble WS (2015) Machine learning applications in genetics and genomics. Nature reviews. Genetics 16(6):321–332. https://doi.org/10.1038/nrg3920

Lin YT, Whitman WB, Coleman DC, Chiu CY (2018) Effects of reforestation on the structure and diversity of bacterial communities in subtropical low mountain forest soils. Front Microbiol 9:1–10. https://doi.org/10.3389/fmicb.2018.01968

Liu M, Wang Z, Li S, Lu X, Wang X, Han X (2017) Changes in specific leaf area of dominant plants in temperate grasslands along a 2500-km transect in northern China. Sci Rep 7:10780. https://doi.org/10.1038/s41598-017-11133-z

Livingston A, Cameron AD, Petty JA, Le SL (2004) Effect of growth rate on the properties of genetically improved Sitka spruce wood. Forestry: An International Journal of Forest Research 77(4):325–334. https://doi.org/10.1093/forestry/77.4.325

Loo J, Souvannavong O, Dawson IK (2014) Seeing the trees as well as the forest: the importance of managing forest genetic resources. For Ecol Manag 333:1–8. https://doi.org/10.1016/j.foreco.2014.08.014

Marco de Lima B, Cappa EP, Silva-Junior OB, Garcia C, Mansfield SD, Grattapaglia D (2019) Quantitative genetic parameters for growth and wood properties in *Eucalyptus* "urograndis" hybrid using near-infrared phenotyping and genome-wide SNP-based relationships. PLoS One 14(6):e0218747. https://doi.org/10.1371/journal.pone.0218747

May A, Hazelhurst S, Li Y, Norris SA, Govind N, Tikly M, Hon C, Johnson KJ, Hartmann N, Staedtler F, Ramsay M (2013) Genetic diversity in black South Africans from Soweto. BMC Genomics 14(644):1–12. https://doi.org/10.1186/1471-2164-14-644

Méndez-González ID, Jardón-Barbolla L, Jaramillo-Correa JP (2017) Differential landscape effects on the fine-scale genetic structure of populations of a montane conifer from central Mexico. Tree Genet Genomes 13(30). https://doi.org/10.1007/s11295-017-1112-5

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics. 157(4):1819–1829

Minamikawa MF, Takada N, Terakami S, Saito T, Onogi A, Kajiya-Kanegae H et al (2018) Genome-wide association study and genomic prediction using parental and breeding populations of Japanese pear (*Pyrus pyrifolia* Nakai) Sci Rep 8:11994. https://doi.org/10.1038/s41598-018-30154-w

Mitchell-Olds T, Willis JH, Goldstein DB (2007) Which evolutionary processes influence natural genetic variation for phenotypic traits? Nat Rev Genet 8:845–856. https://doi.org/10.1038/nrg2207

Money DK, Gardner Z, Migicovsky H, Schwaninger G, Zhong Y et al (2015) LinkImpute: fast and accurate genotype imputation for nonmodel organisms. G3(Bethesda) 5:2383-2390. https://doi.org/10.1534/g3.115.021667

Moran E, Lauder J, Musser C, Stathos A, Shu M (2017) The genetics of drought tolerance in conifers. New Phytol 216:1034–1048. https://doi.org/10.1111/nph.14774

Müller BSF, Neves LG, de Almeida Filho JE, Resende MFR, Muñoz PR, dos Santos PET et al (2017) Genomic prediction in contrast to a genome-wide association study in explaining heritable variation of complex growth traits in breeding populations of Eucalyptus. BMC Genomics 18(524):1–17. https://doi.org/10.1186/s12864-017-3920-2

Norman A, Taylor J, Edwards J, Kuchel H (2018) Optimising genomic selection in wheat: effect of marker density, population size and population structure on prediction accuracy. G3 (Bethesda) 8(9): 2889–2899. https://doi.org/10.1534/g3.118.200311

Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, 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 JA, Sena J, Svensson T, Talavera-López C, Theißen G, Tuominen H, Vanneste K, Wu ZQ, Zhang B, Zerbe P, Arvestad L, Bhalerao R, Bohlmann J, Bousquet J, Garcia Gil R, Hvidsten TR, de Jong P, MacKay J, Morgante M, Ritland K, Sundberg B, Lee Thompson S, van de Peer Y, Andersson B, Nilsson O, Ingvarsson PK, Lundeberg J, Jansson S (2013) The Norway spruce genome sequence and conifer genome evolution. Nature 497:579–584. https://doi.org/10.1038/nature12211

Ortíz-Bibian MA, Blanco-García A, Lindig-Cisneros RA, Gómez-Romero M, Castellanos-Acuña D, Herrerías-Diego Y et al (2017) Genetic variation in *Abies religiosa* for quantitative traits and delineation of elevational and climatic zoning for maintaining monarch butterfly overwintering sites in Mexico, considering climatic change. Silvae Genetica 66:14–23. https://doi.org/10.1515/sg-2017-0003

Pautasso M (2009) Geographical genetics and the conservation of forest trees. Perspectives in plant ecology, Evolution and Systematics 11(3):157–189. https://doi.org/10.1016/j.ppees.2009.01.003

Pérez-Rodríguez P, de los Campos G (2014) Genome-wide regression and prediction with the BGLR statistical package. Genetics 198: 483–495. https://doi.org/10.1534/genetics.114.164442

Pérez-Rodríguez P, de los Campos G, Crossa J, Gianola D (2010) Genomic-enabled prediction based on molecular markers and pedigree using the bayesian linear regression package in R. Plant Genome 3(2):106–116. https://doi.org/10.3835/plantgenome2010.04.0005

Plomion C, Bartholomé J, Bouffier L, Brendel O, Cochard H, De Miguel M et al (2016) Understanding the genetic bases of adaptation to soil water deficit in trees through the examination of water use efficiency and cavitation resistance: maritime pine as a case study. J Plant Hydraul 3:008. https://doi.org/10.20870/jph.2016.e008

Poland J, Endelman J, Dawson J, Rutkoski J, Wu S, Manes Y, Dreisigacker S, Crossa J, Sánchez-Villeda H, Sorrells M, Jannink JL (2012) Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genome 5:103–113. https://doi.org/10.3835/plantgenome2012.06.0006

Poorter H (2002) Plant growth and carbon economy. Encycl Life Sci https://doi.org/10.1038/npg.els.0003200

Purcell S (2010) Plink-Doc-1.07. Book, 1–293. Available at: papers://55069ee6-504c-4f60-bfa9-053c4dcabb39/paper/p904

Rambolarimanana T, Ramamonjisoa L, Verhaegen D, Leong-Pock-Tsy JM, Jacquin L, Cao-Hamadou TV et al (2018) Performance of multi-trait genomic selection for Eucalyptus robusta breeding program. Tree Genetics and Genomes 14. https://doi.org/10.1007/s11295-018-1286-5

Ratcliffe B, Gamal El-Dien O, Cappa EP, Porth I, Klápště J, Chen C et al (2017) Single-step BLUP with varying genotyping effort in open-pollinated Picea glauca. G3 7:935–942. https://doi.org/10.1534/g3.116.037895

Ratcliffe B, Thistlethwaite FR, El-Dien OG, Cappa E, Porth I, Klápště J et al (2019) Inter- and intra-generation genomic predictions for Douglas-fir growth in unobserved environments bioRxiv, 540765. https://doi.org/10.1101/540765

Redmond MD, Davis TS, Ferrenberg S, Wion AP (2019) Resource allocation trade-offs in a mast-seeding conifer: Piñon pine prioritizes reproduction over defence. AoB PLANTS 11:1–11. https://doi.org/10.1093/aobpla/plz070

Resco de Dios V, Arteaga C, Hedo J, Gil-Pelegrín E, Voltas J (2018) A trade-off between embolism resistance and bark thickness in conifers: are drought and fire adaptations antagonistic? Plant Ecology and Diversity 11:253–258. https://doi.org/10.1080/17550874.2018.1504238

Resende MDV, Resende MFR, Sansaloni CP, Petroli CD, Missiaggia AA, Aguiar AM et al (2012a) Genomic selection for growth and wood quality in Eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees. New Phytol 194:116–128. https://doi.org/10.1111/j.1469-8137.2011.04038.x

Resende JFR, Muñoz P, Resende MDV, Garrick DJ, Fernando RL, Davis JM et al (2012b) Accuracy of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda* L). Genetics 190:1503–1510. https://doi.org/10.1534/genetics.111.137026

Robledo-Arnuncio JJ, Alía R, Gil L (2004) Increased selfing and correlated paternity in a small population of a predominantly outcrossing conifer, *Pinus sylvestris*. Mol Ecol 13(9):2567–2577. https://doi.org/10.1111/j.1365-294X.2004.02251.x

Rogers AR (2014) How population growth affects linkage disequilibrium. Genetics 197:1329–1341. https://doi.org/10.1534/genetics.114.166454

Sáenz-Romero C, Rehfeldt GE, Soto-Correa JC, Aguilar-Aguilar S, Zamarripa-Morales V, López-Upton J (2012) Altitudinal genetic variation among Pinus pseudostrobus populations from Michoacán, México: two location shadehouse test results. Rev Fitotec Mex 35(2):111–121

Sáenz-Romero C, Lindig-Cisneros RA, Joyce DG, Beaulieu J, Bradley JS, Jaquish BC (2016) Assisted migration of forest populations for adapting trees to climate change. Rev Chapingo, Ser Ciencias For y del Ambient 22:303–323. https://doi.org/10.5154/r.rchscfa.2014.10.052

Schuster R, Oberhuber W (2013) Drought sensitivity of three co-occurring conifers within a dry inner Alpine environment. Trees. Trees (Berl West) 27(1):61–69. https://doi.org/10.1007/s00468-012-0768-6

Sedlacek J, Cortés AJ, Wheeler J, Bossdorf O, Hoch G, Klápště J, Lexer C, Rixen C, Wipf S, Karrenberg S, Kleunen M (2016) Evolutionary potential in the Alpine: trait heritabilities and performance variation of the dwarf willow from different elevations and microhabitats. Ecol Evol 6(12):3940–3952

Soltys-Kalina D, Plich J, Strzelczyk-Żyta D, Śliwka J, Marczewski W (2016) The effect of drought stress on the leaf relative water content and tuber yield of a half-sib family of 'katahdin'-derived potato cultivars. Breed Sci 66:328–331. https://doi.org/10.1270/jsbbs.66.328

Sork VL (2018) Genomic studies of local adaptation in natural plant populations. J Hered 109:3–15. https://doi.org/10.1093/jhered/esx091

Sousa TV, Caixeta ET, Alkimim ER, Bertrand B (2019) Early selection enabled by the implementation of genomic selection in *Coffea arabica* breeding. Front Plant Sci 9:1–12. https://doi.org/10.3389/fpls.2018.01934

Suontama M, Klápště J, Telfer E, Graham N, Stovold T, Low C, McKinley R, Dungey H (2019) Efficiency of genomic prediction across two *Eucalyptus nitens* seed orchards with different selection histories. Heredity (Edinb) 122:370–379. https://doi.org/10.1038/s41437-018-0119-5

Tan B, Grattapaglia D, Martins GS, Ferreira KZ, Sundberg B, Ingvarsson PK (2017) Evaluating the accuracy of genomic prediction of growth and wood traits in two *Eucalyptus* species and their F1 hybrids.

BMC Plant Biol 17:1–15. https://doi.org/10.1186/s12870-017-1059-6

Telfer E, Graham N, Stanbra L, Manley T, Wilcox P (2013) Extraction of high purity genomic DNA from pine for use in a high-throughput genotyping platform. N Z J For Sci 43:1–8. https://doi.org/10.1186/1179-5395-43-3

The R Core Team (2019) R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. http://www.R-project.org/

Thistlethwaite FR, Ratcliffe B, Klápště J, Porth I, Chen C, Stoehr MU, el-Kassaby YA (2017) Genomic prediction accuracies in space and time for height and wood density of Douglas-fir using exome capture as the genotyping platform. BMC Genomics 18:1–16. https://doi.org/10.1186/s12864-017-4258-5

Thistlethwaite FR, Ratcliffe B, Klápště J, Porth I, Chen C, Stoehr MU, el-Kassaby YA (2019) Genomic selection of juvenile height across a single-generational gap in Douglas-fir. Heredity (Edinb) 122:848–863. https://doi.org/10.1038/s41437-018-0172-0

Thistlethwaite FR, Gamal El-Dien O, Ratcliffe B, Klápště J, Porth I, Chen C et al (2020) Linkage disequilibrium vs. pedigree: genomic selection prediction accuracy in conifer species. PLoS One 15(6): e0232201. https://doi.org/10.1371/journal.pone.0232201

Thomson MJ (2014) High-throughput SNP genotyping to accelerate crop improvement. Plant Breeding and Biotechnology 2(3):195–212. https://doi.org/10.9787/PBB.2014.2.3.195

Urza AK, Sibold JS (2013) Nondestructive aging of postfire seedlings for four conifer species in northwestern Montana. West J Appl For 28: 22–29. https://doi.org/10.5849/wjaf.11-014

Van Eenennaam AL, Weigel KA, Young AE, Cleveland MA, Dekkers JCM (2014) Applied animal genomics :results from the field. Annu Rev Anim Biosci 2:105–139. https://doi.org/10.1146/annurev-animal-022513-114119

VanRaden PM (2008) Efficient methods to compute genomic predictions. J Dairy Sci 91:4414–4423. https://doi.org/10.3168/jds.2007-0980

Wachowiak W, Perry A, Donnelly K, Cavers S (2018) Early phenology and growth trait variation in closely related European pine species. Ecol Evol 8(1):655–666. https://doi.org/10.1002/ece3.3690

Wang T, Hamann A, Yanchuk A, O'neill GA, Aitken SN (2006) Use of response functions in selecting lodgepole pine populations for future climates. Glob Chang Biol 12(12):2404–2416. https://doi.org/10.1111/j.1365-2486.2006.01271.x

Wang T, Hamann A, Spittlehouse DL, Murdock TQ (2012) ClimateWNA-high-resolution spatial climate data for western North America. J Appl Meteorol. Climatol. 51(1):16–29. https://doi.org/10.1175/JAMC-D-11-043.1

Wang J, Santiago E, Caballero A (2016) Prediction and estimation of effective population size. Heredity (Edinb) 117:193–206. https://doi.org/10.1038/hdy.2016.43

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370. https://doi.org/10.2307/2408641

Wickham H. (2010) ggplot2-Book09h Wickham. https://doi.org/10.1007/978-0-387-98141-3

Williamson GB, Wiemann MC (2010) Measuring wood specific gravity. Am J Bot 97:519–524. https://doi.org/10.3732/ajb.0900243

Zapata-Valenzuela J, Isik F, Maltecca C, Wegrzyn J, Neale D, McKeand S, Whetten R (2012) SNP markers trace familial linkages in a cloned population of *Pinus taeda* - prospects for genomic selection. Tree Genetics and Genomes 8:1307–1318

Zapata-Valenzuela J, Whetten RW, Neale D, McKeand S, Isik F (2013) Genomic estimated breeding values using genomic relationship matrices in a cloned population of loblolly pine. G3 3(5):909–916. https://doi.org/10.1534/g3.113.005975

Zhang A, Wang H, Beyene Y, Semagn K, Liu Y, Cao S, Cui Z, Ruan Y, Burgueño J, San Vicente F, Olsen M, Prasanna BM, Crossa J, Yu H, Zhang X (2017) Effect of trait heritability, training population size and marker density on genomic prediction accuracy estimation in 22 bi-parental tropical maize populations. Front Plant Sci 8:1–12. https://doi.org/10.3389/fpls.2017.01916

Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics 28(24):3326–3328. https://doi.org/10.1093/bioinformatics/bts606

Zhou L, Chen Z, Olsson L, Grahn T, Karlsson B, Wu HX, Lundqvist SO, García-Gil MR (2020) Effect of number of annual rings and tree ages on genomic predictive ability for solid wood properties of Norway spruce. BMC Genomics 21:323. https://doi.org/10.1186/s12864-020-6737-3

SUPPLEMENTARY TABLES

Table S1. Geographical coordinates for the 213 individuals of *Abies religiosa* used for GP-modelling.

Individuals						
Origin	**Experimental site**	**Group**	**No. of Individual**	**Age**	**Latitude**	**Longitude**
Reforestation	Rincón de Guadalupe	RF	1	Nine	19.2577	-99.9569
Reforestation	Rincón de Guadalupe	RF	2	Seven	19.2570	-99.9566
Reforestation	Rincón de Guadalupe	RF	3	Nine	19.2572	-99.9565
Reforestation	Rincón de Guadalupe	RF	4	Nine	19.2570	-99.9566
Reforestation	Rincón de Guadalupe	RF	5	Eight	19.2570	-99.9565
Reforestation	Rincón de Guadalupe	RF	6	Eight	19.2569	-99.9565
Reforestation	Rincón de Guadalupe	RF	7	Eight	19.2567	-99.9566
Reforestation	Rincón de Guadalupe	RF	8	Eight	19.2568	-99.9564
Reforestation	Rincón de Guadalupe	RF	9	Eight	19.2568	-99.9566
Reforestation	Rincón de Guadalupe	RF	10	Eight	19.2567	-99.9565
Reforestation	Rincón de Guadalupe	RF	11	Eight	19.2566	-99.9566
Reforestation	Rincón de Guadalupe	RF	12	Eight	19.2565	-99.9564
Reforestation	Rincón de Guadalupe	RF	13	Eight	19.2564	-99.9564
Reforestation	Rincón de Guadalupe	RF	14	Eight	19.2563	-99.9563
Reforestation	Rincón de Guadalupe	RF	15	Eight	19.2562	-99.9560
Reforestation	Rincón de Guadalupe	RF	16	Eight	19.2563	-99.9561
Reforestation	Rincón de Guadalupe	RF	17	Eight	19.2563	-99.9560
Reforestation	Rincón de Guadalupe	RF	18	Eight	19.2563	-99.9554
Reforestation	Rincón de Guadalupe	RF	19	Eight	19.2563	-99.9561
Reforestation	Rincón de Guadalupe	RF	20	Eight	19.2565	-99.9561
Reforestation	Rincón de Guadalupe	RF	21	Seven	19.2566	-99.9562
Reforestation	Rincón de Guadalupe	RF	22	Seven	19.2571	-99.9563
Reforestation	Rincón de Guadalupe	RF	23	Seven	19.2575	-99.9568
Reforestation	Rincón de Guadalupe	RF	24	Seven	19.2579	-99.9570
Reforestation	Rincón de Guadalupe	RF	25	Seven	19.2581	-99.9573
Reforestation	Rincón de Guadalupe	RF	26	Seven	19.2576	-99.9566
Reforestation	Rincón de Guadalupe	RF	27	Seven	19.2577	-99.9565
Reforestation	Rincón de Guadalupe	RF	28	Seven	19.2577	-99.9566
Reforestation	Rincón de Guadalupe	RF	29	Seven	19.2580	-99.9566
Reforestation	Rincón de Guadalupe	RF	30	Seven	19.2576	-99.9567
Reforestation	Rincón de Guadalupe	RF	31	Seven	19.2576	-99.9566
Reforestation	Rincón de Guadalupe	RF	32	Seven	19.2575	-99.9564
Reforestation	Rincón de Guadalupe	RF	33	Seven	19.2576	-99.9564
Reforestation	Rincón de Guadalupe	RF	34	Eight	19.2575	-99.9566
Reforestation	Rincón de Guadalupe	RF	35	Seven	19.2574	-99.9565
Reforestation	Rincón de Guadalupe	RF	36	Seven	19.2574	-99.9567

Reforestation	Rincón de Guadalupe	RF	37	Seven	19.2571	-99.9566
Reforestation	Rincón de Guadalupe	RF	38	Seven	19.2573	-99.9566
Reforestation	Rincón de Guadalupe	RF	39	Seven	19.2568	-99.9563
Reforestation	Rincón de Guadalupe	RF	40	Eight	19.2560	-99.9561
Reforestation	Rincón de Guadalupe	RF	41	Seven	19.2567	-99.9565
Reforestation	Rincón de Guadalupe	RF	42	Eight	19.2567	-99.9563
Reforestation	Rincón de Guadalupe	RF	43	Eight	19.2568	-99.9567
Reforestation	Rincón de Guadalupe	RF	44	Eight	19.2567	-99.9563
Reforestation	Rincón de Guadalupe	RF	45	Eight	19.2567	-99.9563
Reforestation	Rincón de Guadalupe	RF	46	Eight	19.2568	-99.9564
Reforestation	Rincón de Guadalupe	RF	47	Eight	19.2568	-99.9565
Reforestation	Rincón de Guadalupe	RF	48	Eight	19.2569	-99.9565
Reforestation	Rincón de Guadalupe	RF	49	Eight	19.2569	-99.9566
Reforestation	Rincón de Guadalupe	RF	50	Seven	19.2574	-99.9570
Reforestation	Rincón de Guadalupe	RF	51	Seven	19.2581	-99.9572
Reforestation	Rincón de Guadalupe	RF	52	Seven	19.2583	-99.9577
Reforestation	Rincón de Guadalupe	RF	53	Seven	19.2583	-99.9575
Reforestation	Rincón de Guadalupe	RF	54	Seven	19.2583	-99.9575
Reforestation	Rincón de Guadalupe	RF	55	Seven	19.2583	-99.9576
Reforestation	Rincón de Guadalupe	RF	56	Seven	19.2582	-99.9575
Reforestation	Rincón de Guadalupe	RF	57	Seven	-	-
Reforestation	Rincón de Guadalupe	RF	58	Seven	19.2575	-99.9564
Reforestation	Rincón de Guadalupe	RF	59	Eight	19.2563	-99.9565
Reforestation	Rincón de Guadalupe	RF	60	Seven	19.2566	-99.9562
Reforestation	Rincón de Guadalupe	RF	61	Seven	19.2568	-99.9559
Reforestation	Rincón de Guadalupe	RF	62	Seven	19.2567	-99.9558
Reforestation	Rincón de Guadalupe	RF	63	Eight	19.2567	-99.9559
Reforestation	Rincón de Guadalupe	RF	64	Seven	19.2561	-99.9564
Reforestation	Rincón de Guadalupe	RF	65	Eight	19.2564	-99.9564
Reforestation	Rincón de Guadalupe	RF	66	Eight	19.2570	-99.9567
Reforestation	Rincón de Guadalupe	RF	67	Seven	19.2570	-99.9567
Reforestation	Rincón de Guadalupe	RF	68	Seven	19.2572	-99.9556
Reforestation	Rincón de Guadalupe	RF	69	Seven	19.2576	-99.9566
Reforestation	Rincón de Guadalupe	RF	70	Seven	19.2578	-99.9565
Reforestation	Rincón de Guadalupe	RF	71	Seven	19.2571	-99.9563
Reforestation	Rincón de Guadalupe	RF	72	Nine	19.2580	-99.9566
Reforestation	Rincón de Guadalupe	RF	73	Nine	19.2581	-99.9565
Reforestation	Rincón de Guadalupe	RF	74	Nine	19.2578	-99.9562
Reforestation	Rincón de Guadalupe	RF	75	Seven	-	-
Reforestation	Rincón de Guadalupe	RF	76	Nine	19.2572	-99.9565
Reforestation	Rincón de Guadalupe	RF	77	Seven	-	-
Reforestation	Rincón de Guadalupe	RF	78	Seven	19.2560	-99.9567
Reforestation	Rincón de Guadalupe	RF	79	Seven	-	-
Reforestation	Rincón de Guadalupe	RF	80	Nine	19.2578	-99.9560

Reforestation	Rincón de Guadalupe	RF	81	Nine	19.2563	-99.9560
Natural regeneration	Rincón de Guadalupe	NR	1	Eight	19.2576	-99.9567
Natural regeneration	Rincón de Guadalupe	NR	2	Eleven	19.2572	-99.9564
Natural regeneration	Rincón de Guadalupe	NR	3	Twelve	19.2571	-99.9564
Natural regeneration	Rincón de Guadalupe	NR	4	Nine	19.2571	-99.9568
Natural regeneration	Rincón de Guadalupe	NR	5	Nine	19.2574	-99.9569
Natural regeneration	Rincón de Guadalupe	NR	6	Ten	19.2573	-99.9568
Natural regeneration	Rincón de Guadalupe	NR	7	Eight	19.2577	-99.9571
Natural regeneration	Rincón de Guadalupe	NR	8	Seven	19.2579	-99.9573
Natural regeneration	Rincón de Guadalupe	NR	9	Eleven	19.2574	-99.9567
Natural regeneration	Rincón de Guadalupe	NR	10	Ten	19.2565	-99.9563
Natural regeneration	Rincón de Guadalupe	NR	11	Twelve	19.2566	-99.9564
Natural regeneration	Rincón de Guadalupe	NR	12	Nine	19.2566	-99.9563
Natural regeneration	Rincón de Guadalupe	NR	13	Eight	19.2561	-99.9558
Natural regeneration	Rincón de Guadalupe	NR	14	Seven	19.2575	-99.9564
Natural regeneration	Rincón de Guadalupe	NR	15	Nine	19.2580	-99.9571
Natural regeneration	Rincón de Guadalupe	NR	16	Nine	19.2561	-99.9558
Natural regeneration	Rincón de Guadalupe	NR	17	Eight	19.2567	-99.9564
Natural regeneration	Rincón de Guadalupe	NR	18	Eight	-	-
Natural regeneration	Rincón de Guadalupe	NR	19	Eight	19.2563	-99.9559
Natural regeneration	Rincón de Guadalupe	NR	20	Ten	19.2575	-99.9570
Natural regeneration	Rincón de Guadalupe	NR	21	Nine	-	-
Natural regeneration	Rincón de Guadalupe	NR	22	Nine	19.2573	-99.9567
Natural regeneration	Rincón de Guadalupe	NR	23	Seven	19.2564	-99.9562
Natural regeneration	Rincón de Guadalupe	NR	24	Ten	19.2581	-99.9570
Natural regeneration	Rincón de Guadalupe	NR	25	Ten	19.2575	-99.9570
Natural regeneration	Rincón de Guadalupe	NR	26	Seven	19.2577	-99.9571
Natural regeneration	Rincón de Guadalupe	NR	27	Seven	19.2577	-99.9571
Natural regeneration	Rincón de Guadalupe	NR	28	Nine	-	-
Natural regeneration	Rincón de Guadalupe	NR	29	Nine	19.2574	-99.9569
Natural regeneration	Rincón de Guadalupe	NR	30	Seven	19.2567	-99.9559
Natural regeneration	Rincón de Guadalupe	NR	31	Eight	19.2564	-99.9558
Natural regeneration	Rincón de Guadalupe	NR	32	Eight	19.2574	-99.9566
Natural regeneration	Rincón de Guadalupe	NR	33	Seven	19.2573	-99.9567
Natural regeneration	Rincón de Guadalupe	NR	34	Nine	19.2575	-99.9566
Natural regeneration	Rincón de Guadalupe	NR	35	Nine	-	-
Natural regeneration	Rincón de Guadalupe	NR	36	Nine	-	-
Natural regeneration	Rincón de Guadalupe	NR	37	Nine	19.2574	-99.9564
Natural regeneration	Rincón de Guadalupe	NR	38	Eight	-	-
Natural regeneration	Rincón de Guadalupe	NR	39	Ten	-	-
Natural regeneration	Rincón de Guadalupe	NR	40	Nine	19.2554	-99.9555
Natural regeneration	Rincón de Guadalupe	NR	41	Nine	-	-
Natural regeneration	Rincón de Guadalupe	NR	42	Nine	-	-
Natural regeneration	Rincón de Guadalupe	NR	43	Nine	-	-

Natural regeneration	Rincón de Guadalupe	NR	44	Seven	-	-
Natural regeneration	Rincón de Guadalupe	NR	45	Ten	-	-
Natural regeneration	Rincón de Guadalupe	NR	46	Ten	-	-
Natural regeneration	Rincón de Guadalupe	NR	47	Eight	19.2557	-99.9554
Natural regeneration	Rincón de Guadalupe	NR	48	Eight	19.2571	-99.9563
Natural regeneration	Rincón de Guadalupe	NR	49	Eight	19.2566	-99.9559
Natural regeneration	Rincón de Guadalupe	NR	50	Eight	19.2558	-99.9574
Natural regeneration	Rincón de Guadalupe	NR	51	Eight	19.2575	-99.9574
Natural regeneration	Rincón de Guadalupe	NR	52	Seven	-	-
Natural regeneration	San Bartolo	SB	1	Eleven	19.2352	-99.9210
Natural regeneration	San Bartolo	SB	2	Ten	19.2353	-99.9210
Natural regeneration	San Bartolo	SB	3	Nine	19.2355	-99.9211
Natural regeneration	San Bartolo	SB	4	Nine	19.2349	-99.9214
Natural regeneration	San Bartolo	SB	5	Nine	19.2349	-99.9213
Natural regeneration	San Bartolo	SB	6	Nine	19.2348	-99.9214
Natural regeneration	San Bartolo	SB	7	Eight	19.2348	-99.9215
Natural regeneration	San Bartolo	SB	8	Eight	19.2347	-99.9216
Natural regeneration	San Bartolo	SB	9	Ten	19.2346	-99.9214
Natural regeneration	San Bartolo	SB	10	Nine	19.2347	-99.9216
Natural regeneration	San Bartolo	SB	11	Ten	19.2340	-99.9229
Natural regeneration	San Bartolo	SB	12	Eleven	19.2345	99.9212
Natural regeneration	San Bartolo	SB	13	Twelve	-	-
Natural regeneration	San Bartolo	SB	14	Ten	19.2338	99.9212
Natural regeneration	San Bartolo	SB	15	Ten	19.2336	99.9216
Natural regeneration	San Bartolo	SB	16	Ten	19.2334	99.9212
Natural regeneration	San Bartolo	SB	17	Eleven	19.2333	99.9211
Natural regeneration	San Bartolo	SB	18	Twelve	-	-
Natural regeneration	San Bartolo	SB	19	Nine	-	-
Natural regeneration	San Bartolo	SB	20	Twelve	19.2342	99.9212
Natural regeneration	San Bartolo	SB	21	Seven	19.2347	-99.9217
Natural regeneration	San Bartolo	SB	22	Nine	19.2346	-99.9221
Natural regeneration	San Bartolo	SB	23	Eleven	19.2345	-99.9223
Natural regeneration	San Bartolo	SB	24	Nine	19.2342	-99.9225
Natural regeneration	San Bartolo	SB	25	Twelve	19.2337	-99.9228
Natural regeneration	San Bartolo	SB	26	Seven	19.2332	-99.9231
Natural regeneration	San Bartolo	SB	27	Seven	19.2331	-99.9242
Natural regeneration	San Bartolo	SB	28	Nine	19.2328	-99.9241
Natural regeneration	San Bartolo	SB	29	Eight	19.2321	-99.9242
Natural regeneration	San Bartolo	SB	30	Ten	19.2315	-99.9243
Natural regeneration	San Bartolo	SB	31	Eleven	19.2324	99.9239
Natural regeneration	San Bartolo	SB	32	Seven	19.2347	-99.9221
Natural regeneration	San Bartolo	SB	33	Twelve	19.2347	-99.9217
Natural regeneration	San Bartolo	SB	34	Six	-	-
Natural regeneration	San Bartolo	SB	35	Nine	-	-

Natural regeneration	San Bartolo	SB	36	Ten	-	-
Natural regeneration	San Bartolo	SB	37	Seven	19.2347	-99.9221
Natural regeneration	San Bartolo	SB	38	Seven	19.2316	-99.9241
Natural regeneration	San Bartolo	SB	39	Ten	-	-
Natural regeneration	San Bartolo	SB	40	Twelve	-	-
Natural regeneration	San Bartolo	SB	41	Nine	-	-
Natural regeneration	San Bartolo	SB	42	Nine	-	-
Natural regeneration	San Bartolo	SB	43	Nine	-	-
Natural regeneration	San Bartolo	SB	44	Six	19.2268	-99.9219
Natural regeneration	San Bartolo	SB	45	Seven	19.2268	-99.9219
Natural regeneration	San Bartolo	SB	46	Seven	19.2269	-99.9217
Natural regeneration	San Bartolo	SB	47	Seven	19.2270	-99.9214
Natural regeneration	San Bartolo	SB	48	Ten	19.2273	-99.9215
Natural regeneration	San Bartolo	SB	49	Seven	19.2274	-99.9215
Natural regeneration	San Bartolo	SB	50	Seven	19.2273	-99.9215
Natural regeneration	San Bartolo	SB	51	Eight	19.2274	-99.9215
Natural regeneration	San Bartolo	SB	52	Seven	19.2276	-99.9217
Natural regeneration	San Bartolo	SB	53	Six	19.2275	-99.9218
Natural regeneration	San Bartolo	SB	54	Nine	19.2277	-99.9221
Natural regeneration	San Bartolo	SB	55	Nine	19.2277	-99.9222
Natural regeneration	San Bartolo	SB	56	Nine	19.2280	-99.9221
Natural regeneration	San Bartolo	SB	57	Nine	19.2281	-99.9219
Natural regeneration	San Bartolo	SB	58	Nine	19.2280	-99.9218
Natural regeneration	San Bartolo	SB	59	Eight	19.2278	-99.9219
Natural regeneration	San Bartolo	SB	60	Nine	19.2277	-99.9220
Natural regeneration	San Bartolo	SB	61	Nine	19.2276	-99.9219
Natural regeneration	San Bartolo	SB	62	Eight	19.2277	-99.9219
Natural regeneration	San Bartolo	SB	63	Eight	19.2277	-99.9217
Natural regeneration	San Bartolo	SB	64	Seven	19.2275	-99.9215
Natural regeneration	San Bartolo	SB	65	Seven	19.2276	-99.9229
Natural regeneration	San Bartolo	SB	66	Six	19.2210	-99.9278
Natural regeneration	San Bartolo	SB	67	Seven	19.2290	-99.9217
Natural regeneration	San Bartolo	SB	68	Six	19.2220	-99.9213
Natural regeneration	San Bartolo	SB	69	Six	19.2324	-99.9239
Natural regeneration	San Bartolo	SB	70	Seven	19.2282	-99.9215
Natural regeneration	San Bartolo	SB	71	Seven	19.2282	-99.9216
Natural regeneration	San Bartolo	SB	72	Seven	19.2282	-99.9216
Natural regeneration	San Bartolo	SB	73	Seven	19.2282	-99.9219
Natural regeneration	San Bartolo	SB	74	Seven	19.2282	-99.9219
Natural regeneration	San Bartolo	SB	75	Seven	19.2281	-99.9219
Natural regeneration	San Bartolo	SB	76	Seven	19.2281	-99.9218
Natural regeneration	San Bartolo	SB	77	Six	19.2280	-99.9222
Natural regeneration	San Bartolo	SB	78	Seven	19.2276	-99.9222
Natural regeneration	San Bartolo	SB	79	Seven	-	-

Natural regeneration	San Bartolo	SB	80	Seven	-	-

Table S2. Five cross-validation schemes in the four pre-established 'across-groups' statistical models to test for GP-predictability in *Abies religiosa*. Predictive models were developed between a training group (G_{TRN}) and validated in a testing group (G_{TST}). Note that the size and composition of training/testing sets ratio varies for each model.

Model ($G_{TRN} \rightarrow G_{TST}$)	**Training set size**				
	TRN = 0%	**TRN = 10%**	**TRN = 20%**	**TRN = 30%**	**TRN = 40%**
NR → RF	51/77	59/69	66/62	74/54	82/46
SB → RF	73/77	81/69	88/62	96/54	104/46
NR → SB	51/73	58/66	66/58	73/51	80/44
SB → NR	73/51	78/46	83/41	88/36	93/31

Table S3. Pairwise-F_{ST} (Weir and Cockerham, 1984) between groups of individuals of *Abies religiosa* in two test trials from central Mexico.

F_{ST}-pairwise	**SB**	**NR**	**RF**
SB	0.0000	0.0108	0.0103
NR	0.0108	0.0000	0.0114
RF	0.0103	0.0114	0.0000

Table S4. Pairwise-F_{ST} (Weir and Cockerham, 1984) between age classes of individuals of *Abies religiosa* in two test trials from central Mexico.

F_{ST}-pairwise	**Six**	**Seven**	**Eight**	**Nine**	**Ten**	**Eleven**	**Twelve**
Six	0.0000	0.0047	0.0090	0.0026	0.0090	0.0101	0.0145
Seven	0.0047	0.0000	0.0008	0.0023	0.0036	0.0077	0.0076
Eight	0.0090	0.0008	0.0000	0.0025	0.0036	0.0043	0.0071
Nine	0.0026	0.0023	0.0025	0.0000	0.0011	0.0020	0.0052
Ten	0.0090	0.0036	0.0036	0.0011	0.0000	0.0026	0.0056
Eleven	0.0101	0.0077	0.0043	0.0020	0.0026	0.0000	0.0101
Twelve	0.0145	0.0076	0.0071	0.0052	0.0056	0.0101	0.0000

Table S5. Mean and standard error of predictive ability with two predictive methods for ten traits in *Abies religiosa*. The model with the lowest mean observed standard error is indicated in bold.

Phenotypic Trait	***BRR***	***RKHS***
	r_y	r_y
Growth traits		
Total height (TH)	0.402 (0.02)[a1,2]	0.426 (0.04)[a]
Stem diameter (SD)	0.490 (0.02)[a]	0.524 (0.03)[b]
Above-ground biomass (AGB)	0.443 (0.03)[a]	0.452 (0.03)[a]
Crown ratio (CR)	0.383 (0.03)[a]	0.358 (0.05)[a]
2016's Growth (G_{2016})	0.541 (0.02)[a]	0.570 (0.02)[b]
Wood density (WD)	0.322 (0.01)[a]	0.309 (0.01)[a]
Physiological traits		
Water potential (Ψ)	0.434 (0.03)[a]	0.412 (0.03)[a]
Relative water content (RWC)	0.229 (0.03)[a]	0.248 (0.03)[a]
Specific leaf area (SLA)	0.286 (0.03)[a]	0.259 (0.03)[a]
Relative growth rate (RGR)	0.090 (0.04)[a]	0.132 (0.03)[b]
Average	**0.362 (0.01)**[a]	0.369 (0.02)[a]

[1] Mean and standard for error of predictive ability of each method were calculated taking a model trained with 90% of individuals and tested with 10% of individuals for whole sample (combined model);

[2] Different alphabetic letters indicate significant differences (p<0.05) according to Tukey's HSD, between the methods for each trait after one-way ANOVA.

Table S6. Mean and standard error for heritability values and estimated predictive ability for the two types of traits estimated in 201 *Abies religiosa* individuals.

	Growth traits	***Physiological traits***
h^2	0.287 (0.01)[a1]	0.196 (0.01)[b]
r_y	0.423 (0.03)[a]	0.269 (0.01)[b]

[1]Different letters denote significant differences (p < 0.05) after a one-way ANOVA.

Table S7. Mean and standard errors for predictive abilities of ten phenotypic traits in three groups of *Abies religiosa* individuals.

Phenotypic Trait	***SB***	***NR***	***RF***
	r_y	r_y	r_y
Growth traits			
Total height (TH)	0.46 (0.02)^{a1,2}	0.35 (0.02)^{b}	0.29 (0.03)^{c}
Stem diameter (SD)	0.42 (0.03)^{a}	0.45 (0.02)^{a}	0.38 (0.02)^{b}
Above-ground biomass (AGB)	0.35 (0.03)^{a}	0.42 (0.03)^{b}	0.37 (0.02)^{a}
Crown ratio (CR)	0.51 (0.03)^{a}	0.25 (0.03)^{b}	0.19 (0.02)^{b}
2016's Growth (G_{2016})	0.45 (0.03)^{a}	0.58 (0.04)^{b}	0.44 (0.05)^{a}
Wood density (WD)	0.3 (0.04)^{a}	0.23 (0.03)^{a}	0.11 (0.05)^{b}
Physiological traits			
Water potential (Ψ)	0.39 (0.04)^{a}	0.41 (0.03)^{a}	0.37 (0.03)^{a}
Relative water content (RWC)	0.21 (0.05)^{a}	0.14 (0.05)^{b}	0.04 (0.05)^{c}
Specific leaf area (SLA)	0.28 (0.05)^{a}	0.15 0.03)^{b}	0.22 (0.03)^{a}
Relative growth rate (RGR)	-0.1 (0.08)^{a}	0.15 (0.06)^{a}	0.02 (0.08)^{a}
Average	**0.327 (0.01)^{a}**	**0.313 (0.02)^{a}**	**0.243 (0.02)^{b}**

[1]Predictive ability calculated with models trained with 90% of individuals and tested with 10% of trees using the BRR method; Different letters indicate significant differences between groups ($p < 0.05$) according to Tukey's HSD after a one-way ANOVA.

Table S8. Mean and standard errors for predictive abilities estimated for four 'across-groups' models with five training/testing dataset (TRN/TST) ratios, expressed in percentage and numbers of individuals, and calculated with the BRR method.

Model	TRN/TST	TH	SD	AGB	CR	G_{2016}	WD	Ψ	RWC	SLA	RGR
NR → RF	**0% (51/77)**	0.124 (0.03)	0.039 (0.04)	0.032 (0.04)	0.113 (0.04)	0.082 (0.03)	0.152 (0.02)	0.137 (0.03)	0.131 (0.03)	-0.081 (0.03)	0.120 (0.03)
NR → RF	**10% (59/69)**	0.239 (0.05)	0.155 (0.05)	0.201 (0.03)	0.232 (0.03)	0.299 (0.03)	0.227 (0.05)	0.291 (0.04)	0.110 (0.02)	0.071 (0.02)	0.094 (0.03)
NR → RF	**20% (66/62)**	0.084 (0.03)	0.268 (0.04)	0.265 (0.04)	0.300 (0.02)	0.368 (0.05)	0.276 (0.04)	0.346 (0.02)	0.168 (0.04)	0.069 (0.03)	0.118 (0.04)
NR → RF	**30% (74/54)**	0.330 (0.04)	0.309 (0.03)	0.307 (0.03)	0.262 (0.03)	0.418 (0.05)	0.258 (0.03)	0.397 (0.03)	0.163 (0.04)	0.192 (0.04)	0.106 (0.04)
NR → RF	**40% (82/46)**	0.260 (0.06)	0.358 (0.03)	0.340 (0.02)	0.309 (0.04)	0.438 (0.03)	0.225 (0.01)	0.404 (0.02)	0.240 (0.04)	0.182 (0.05)	0.129 (0.03)
SB → RF	**0% (73/77)**	0.134 (0.03)	0.183 (0.01)	0.105 (0.02)	0.043 (0.01)	-0.12 (0.05)	0.143 (0.01)	-0.006 (0.04)	0.018 (0.04)	0.188 (0.03)	0.093 (0.04)
SB → RF	**10% (81/69)**	0.102 (0.03)	0.154 (0.02)	0.124 (0.4)	-0.005 (0.02)	0.02 (0.04)	0.132 (0.01)	-0.054 (0.05)	-0.01 (0.03)	0.167 (0.04)	0.082 (0.03)
SB → RF	**20% (88/62)**	0.181 (0.02)	0.173 (0.03)	0.129 (0.03)	0.125 (0.02)	0.136 (0.02)	0.189 (0.04)	0.106 (0.02)	0.068 (0.03)	0.180 (0.02)	0.075 (0.03)
SB → RF	**30% (96/54)**	0.155 (0.05)	0.234 (0.03)	0.186 (0.04)	0.076 (0.02)	0.226 (0.03)	0.232 (0.03)	0.183 (0.05)	0.123 (0.03)	0.246 (0.03)	0.069 (0.05)
SB → RF	**40% (104/46)**	0.136 (0.02)	0.228 (0.02)	0.143 (0.03)	0.081 (0.01)	0.253 (0.03)	0.226 (0.02)	0.176 (0.03)	0.134 (0.04)	0.219(0.03)	0.104 (0.02)
NR → SB	**0% (51/73)**	0.063 (0.02)	-0.047 (0.07)	-0.025 (0.02)	0.137 (0.03)	0.064 (0.01)	0.073 (0.02)	0.234 (0.01)	0.12 (0.02)	0.158 (0.02)	0.03 (0.02)
NR → SB	**10% (58/66)**	0.195 (0.05)	0.052 (0.09)	0.02 (0.03)	0.179 (0.02)	0.175 (0.04)	0.118 (0.02)	0.253 (0.01)	0.153 (0.01)	0.116 (0.02)	0.037 (0.04)
NR → SB	**20% (66/58)**	0.203 (0.02)	0.193 (0.07)	0.116 (0.04)	0.272 (0.01)	0.182 (0.03)	0.147 (0.01)	0.279 (0.03)	0.161 (0.01)	0.171 (0.04)	0.101 (0.05)
NR → SB	**30% (73/51)**	0.112 (0.08)	0.289 (0.08)	0.207 (0.06)	0.344 (0.04)	0.316 (0.02)	0.181 (0.03)	0.311 (0.02)	0.229 (0.04)	0.303 (0.03)	0.101 (0.03)
NR → SB	**40% (80/44)**	0.212 (0.03)	0.30 (0.07)	0.202 (0.05)	0.338 (0.03)	0.332 (0.04)	0.105 (0.02)	0.332 (0.03)	0.243 (0.04)	0.26 (0.04)	0.138 (0.05)
SB → NR	**0% (73/51)**	0.030 (0.03)	-0.062 (0.03)	-0.210 (0.01)	0.058 (0.03)	0.173 (0.02)	0.094 (0.03)	0.113 (0.03)	0.065 (0.03)	0.252 (0.02)	0.068 (0.02)
SB → NR	**10% (78/46)**	-0.01 (0.03)	-0.083 (0.04)	-0.155 (0.02)	0.103 (0.04)	0.187 (0.01)	-0.049 (0.04)	0.289 (0.05)	0.00 (0.05)	0.232 (0.04)	-0.024 (0.04)
SB → NR	**20% (83/41)**	0.090 (0.02)	-0.075 (0.05)	-0.120 (0.03)	0.081 (0.03)	0.236 (0.05)	0.198 (0.03)	0.350 (0.04)	0.243 (0.03)	0.162 (0.04)	-0.069 (0.05)
SB → NR	**30% (88/36)**	0.003 (0.04)	0.088 (0.03)	-0.053 (0.03)	0.127 (0.04)	0.296 (0.03)	0.211 (0.04)	0.297 (0.05)	0.273 (0.04)	0.258 (0.05)	0.097 (0.05)
SB → NR	**40% (93/31)**	0.160 (0.03)	0.069 (0.06)	0.136 (0.04)	0.136 (0.05)	0.304 (0.04)	0.261 (0.05)	0.331 (0.04)	0.172 (0.03)	0.221 (0.07)	0.056 (0.06)

SUPPLEMENTARY FIGURES

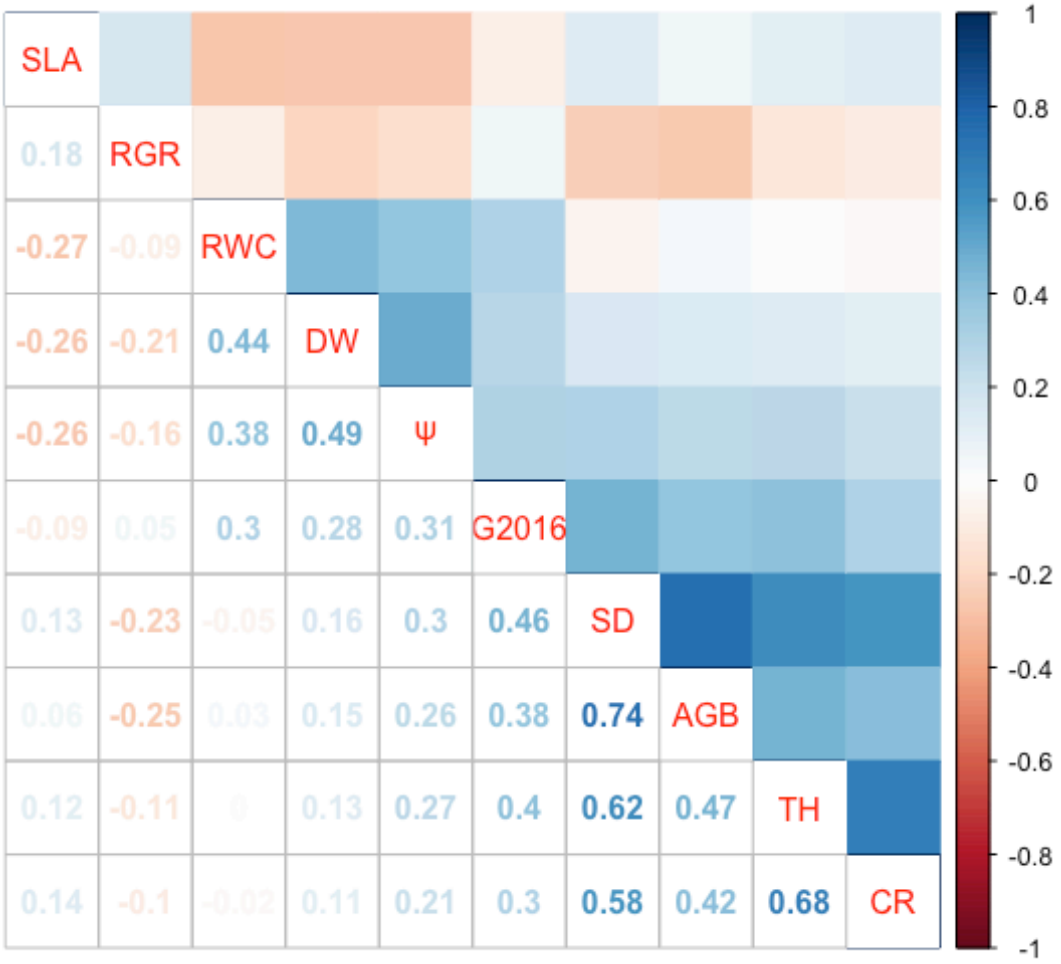

Fig S1. Corplot showing statistically significant relationship between phenotypic traits in *A. religiosa*. Numbers within boxes are correlation coefficients. Red indicates a significantly negative correlation, blue a significantly positive correlation and white non-statistical significance.

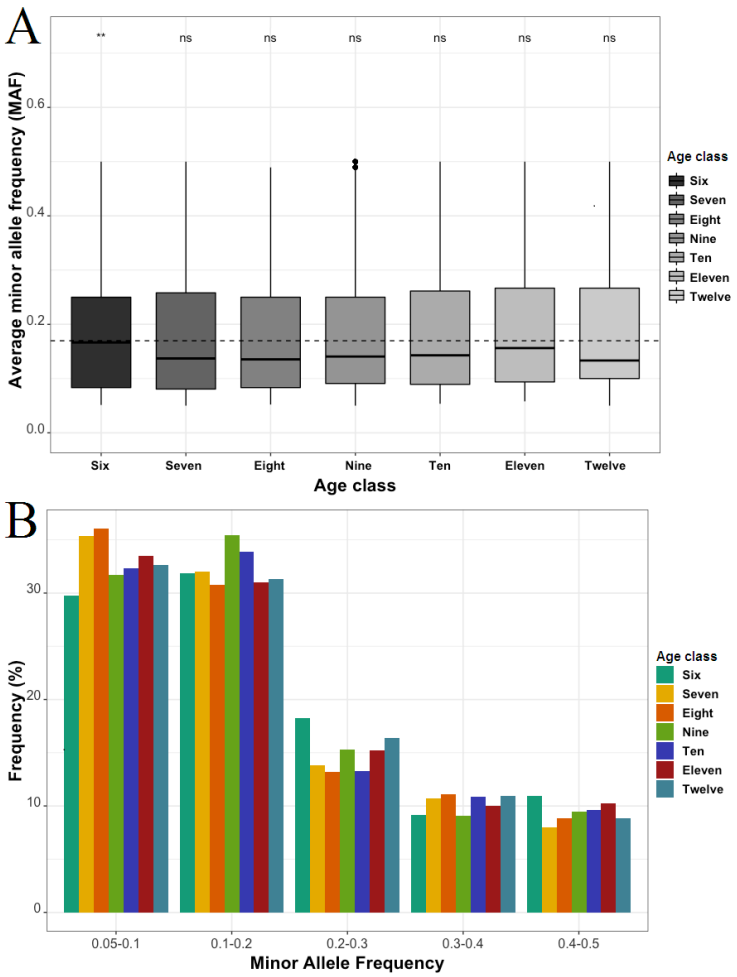

Fig. S2A. Boxplot of average minor allele frequency (MAF) for SNP markers and **S2B** percentage frequency distribution of MAF values in seven age classes.

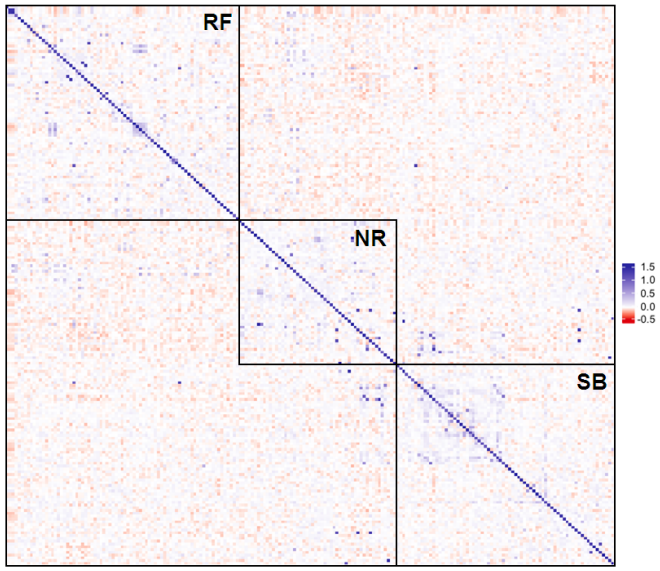

Fig S3. Heatmap of the pairwise genomic relationship matrix (*G*) of Sacred fir trees genotyped with 2286 SNP markers. Heatmap scales show a continuum of realized genetic relationships between individual pairs, from low relationships (red areas corresponding to values below zero), increasing to half-sib relationships (light blue shades, around 0.25), to full-sibs (blue areas, values around 0.5).

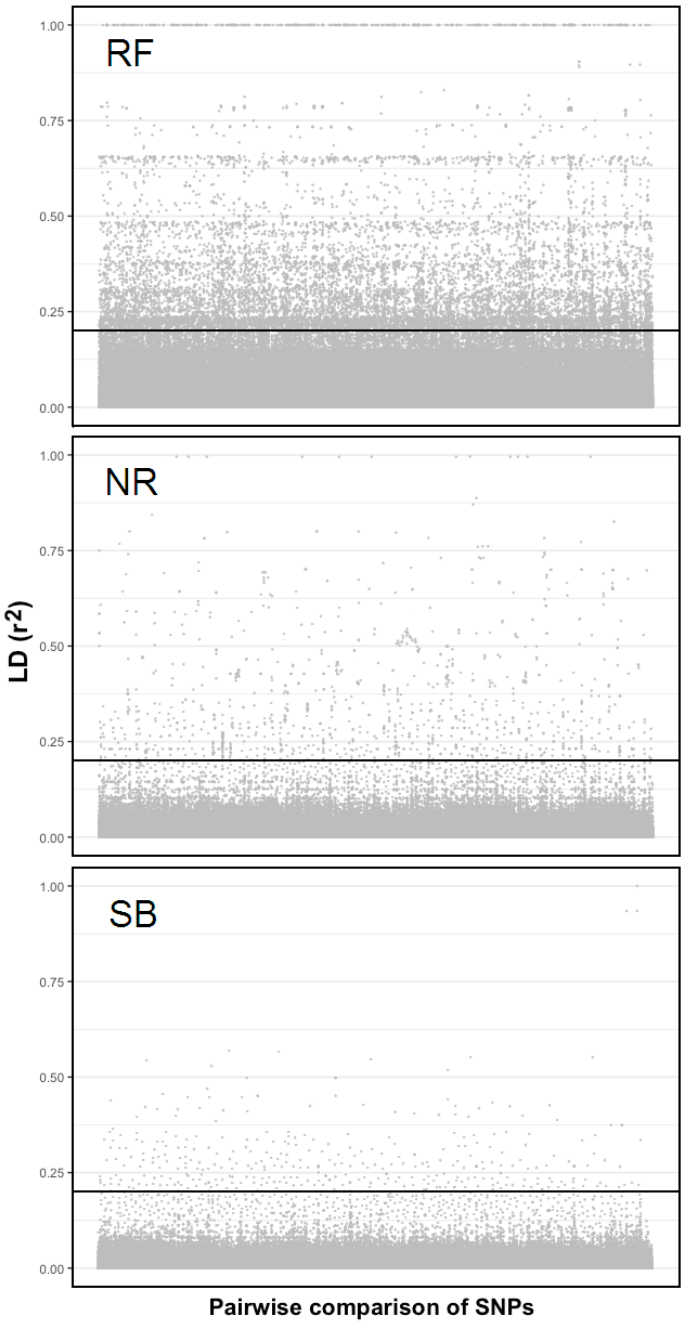

Fig S4. Scatterplot of pairwise linkage disequilibrium between SNPs (LD) (r^2) for three group of *Abies religiosa* individuals.

CAPÍTULO V

DISCUSIÓN Y PERSPECTIVAS

5.1. Discusión general

Comprender cuáles son los componentes de la variación fenotípica que responden a la heterogeneidad ambiental proporciona información para conservar y manejar las poblaciones naturales y favorecer su adaptación y supervivencia en el largo plazo. Una pregunta fundamental en ecología evolutiva es cómo el fenotipo responde a estos componentes y cuáles son los mecanismos involucrados. Por lo tanto, estudiar el impacto ecológico y genético, así como las implicaciones evolutivas de la variación fenotípica, es de gran importancia para el manejo y la conservación de la diversidad biológica. Los estudios que integran la presente tesis permitieron abordar preguntas sobre la configuración de los patrones de variación fenotípica dentro de poblaciones naturales de oyamel (*A. religiosa*). Luego de una introducción general, se analizaron inicialmente algunos modelos poligénicos microgeográficos para caracterizar la diferenciación genética, las presiones de la heterogeneidad del suelo y la interacción G × E sobre la variación cuantitativa en dos grupos de árboles con orígenes diferentes (regeneración natural y reforestación; capítulo II). En el segundo estudio (capítulo III), se revisaron las posibilidades teóricas de utilizar las herramientas de selección y predicción genómica en poblaciones naturales de árboles, mismas que han sido validadas en plantaciones de especies semi-domesticadas como pinos (Zapata-Valenzuela et al., 2013; Isik et al., 2015, 2016) o abetos (Lenz et al., 2017; Chen et al., 2018). En el tercer estudio (capítulo IV), se utilizaron herramientas estadísticas de predicción genómica (PG) para predecir el desempeño fenotípico dentro de un estudio piloto en poblacionales naturales de oyamel.

Desempeño de los modelos poligénicos y de predicción genómica (PG) en poblaciones naturales

En el capítulo II se evaluaron modelos poligénicos para explicar/predecir la variación fenotípica de rasgos fisiológicos (7 rasgos) y de crecimiento (12 rasgos) en individuos jóvenes regenerados naturalmente (pool genético local) y reforestados (pool genético introducido) que coexisten entremezclados en un rodal forestal (“Rincón de Guadalupe”) en el centro de México. Se encontró, que a pesar de que no hay información confiable sobre el origen de las semillas de reforestación, los grupos tienen poca diferenciación (F_{ST} = 0.012) y niveles similares de variación genética (heterocigosidades observadas de 0.235 y 0.219 para NR y RF, respectivamente; p = 0.092). Estos niveles bajos son además similares a los reportados en otras especies forestales a partir de

marcadores obtenidos por secuenciación masiva (Hiraoka et al., 2018; Martins et al., 2018; Collevatti et al., 2019) e indican que los regímenes de manejo antrópico han tenido poco impacto sobre la diversidad genética neta de la población analizada.

La evidente distribución restringida de los grupos poblacionales en el área de estudio puede explicar los valores de variación genética en las poblaciones analizadas (capítulos II y IV), aunado a la compleja interacción entre los distintos atributos de historia de vida de la especie estudiada, como su longevidad, la elevada mortalidad en estados juveniles y la reducida dispersión a larga distancia (Ángeles-Cervantes y López-Mata, 2009). También cabe destacar la influencia de factores ambientales e históricos que pueden haber variado en el tiempo y el espacio (Mastretta-Yanes et al., 2018). Siendo así, la actual distribución relictual del oyamel podría ser resultado de uno o múltiples cuellos de botella que tuvieron un impacto histórico en poblaciones que anteriormente poseían una distribución más amplia (Aguirre-Planter et al., 2000). La diferenciación genética débil (Fig. 2; capítulo IV), los valores bajos de endogamia, y un incremento en los niveles de consanguinidad (capítulos II y IV) indican que todos los individuos analizados provienen de un mismo pool genético ancestral. Siendo que la distancia entre San Bartolo y Rincón de Guadalupe es muy corta (7.4 km), es posible que su separación haya sido reciente. Es interesante que estos valores de reducida diferenciación genética son buenos indicadores de que se pueden desarrollar modelos predictivos partir de datos genómicos (capítulo III).

Se encontró una amplia variación fenotípica en varios de los rasgos morfo-fisiológicos evaluados en los tres grupos de árboles de oyamel (NR, SB y RF). Esta gran variación es recurrente en coníferas mexicanas, incluso en aquellas con distribución restringida (Alfaro et al., 2014; Gernandt and Pérez-De La Rosa, 2014). Los individuos regenerados naturalmente (NR y SB) tuvieron mayor crecimiento y mejor desempeño fisiológico que los introducidos (RF), mientras que los individuos de los dos grupos naturales tuvieron un desempeño fisiológico similar (Tabla 1; capítulo IV). En el capítulo II se observó que algunos de estos fenotipos (p. ej., TH, SD de crecimiento y Ψ y P_{leaf} de fisiología), se correlacionan con el grado de diferenciación genética entre individuos (GRE1, Fig. 2B), tal como se ha reportado en estudios previos (Gömöry et al., 2013; Harter et al., 2015). Los rasgos de crecimiento fueron los que más aportaron a la diferenciación entre grupos ($Model_{G1}$ *vs*. $Model_{G2}$, Tabla 2 y S6), evidenciando que los rasgos fisiológicos pueden estar influenciados por una mayor cantidad de genes (Bouvet

et al., 2020) o estar sujetos a una mayor plasticidad (Mitchell-Odds et at., 2007) que los rasgos de crecimiento analizados (p. ej. altura total, diámetro del tallo y radio del docel).

En concordancia con lo esperado en un modelo poligénico, los modelos de PG mostraron heredabilidades genómicas (h_g^2) y habilidades de predicción (r_y) menores para los rasgos fisiológicos (con la excepción del potencial hídrico), que para los caracteres de crecimiento (Tabla 3 y S6; capítulo IV); evidenciando que los primeros estarían más afectados por la variación ambiental y podrían presentar sesgos asociados al momento de la fenotipificación, dependiendo de las condiciones estacionales/climáticas en las que se midan (Plomion et al., 2016; Lenz et al., 2020). Esto a su vez podría disminuir aún más el poder predictivo de los modelos (capítulo III). Estos rasgos fisiológicos son el resultado de procesos complejos de expresión espacial y temporal, donde interactúan muchos genes (Sasaki, 2008). Aun así, son rasgos importantes para medir el efecto del ambiente sobre el genotipo, por lo que ya se están implementando métodos indirectos (*proxies*) para medir rasgos potencialmente variables y conjeturar de manera indirecta la respuesta fisiológica al ambiente. Por ejemplo, calculando la eficiencia en el uso de agua (*WUE*; Bouvet et al., 2020), la respuesta bioquímica a las plagas (Beaulieu et al., 2020) y la relación de biomasa foliar como medida de la ganancia de carbono (Poorter, 2002). Por consiguiente, es de suma importancia incluir métodos de fenotipificación que permitan estimar estas medidas, sobre todo después de algún evento ambiental extremo (p. ej., sequías o heladas severas) en gran cantidad de plantas y (en lo posible) de forma no invasiva, para ser incluidas en modelos predictivos (ver capítulo III).

En esta tesis, los parámetros utilizados para cuantificar el estatus de las relaciones hídricas foliares (i.e., Ψ y RWC) reflejaron un mayor ingreso de agua en los grupos naturalmente regenerados (NR y SB) que en los reforestados (capítulos II y IV). Lo cual reflejaría una mayor capacidad de crecimiento a largo plazo (al menos en sus primeros años de vida). Dada la distribución geográfica restringida de *A. religiosa*, mantener un adecuado suministro hídrico foliar es importante en los ecosistemas de alta montaña, como el Nevado de Toluca, donde la estación seca es bien definida y la precipitación es variable a lo largo del año (i.e., entre junio y septiembre es más intensa; Figura 1). Se sugiere que sin la capacidad de mantener un buen suministro hídrico, la tasa de supervivencia de estas poblaciones podría ser mucho menor e incluso podría impulsar una mayor dormancia fisiológica.

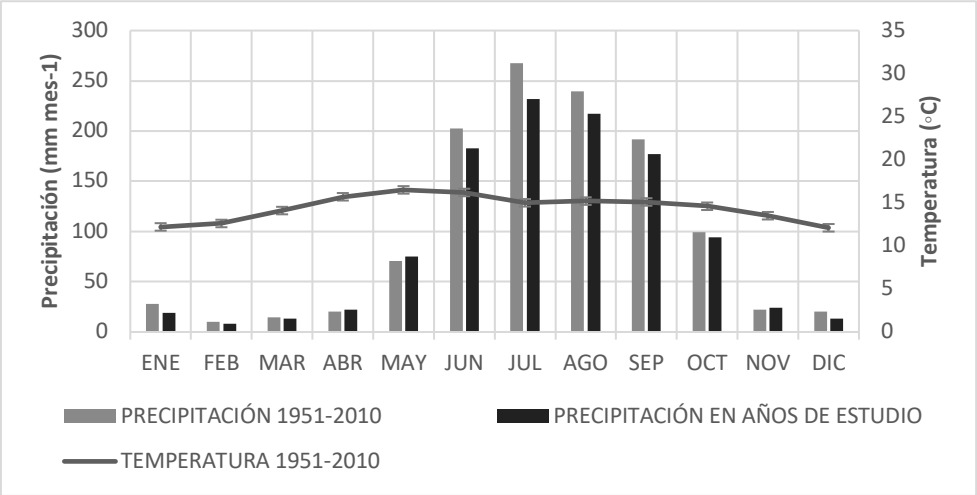

Fig. 1 Climograma desde 1950 hasta el presente del APFFNT. Las barras grises muestran el promedio de la precipitación a lo largo de los años desde 1951 hasta el presente y las barras negras muestran el promedio en los dos años de muestreo (2016 y 2017). La raya representa la variación de temperatura a lo largo del año en cada mes.

Se ha reportado que en las primeras etapas de vida, la sombra natural bajo el dosel es un factor que aumenta la supervivencia de las plántulas de especies arbóreas en bosque templado (incluidas las de *Abies*), ya que esta disminuye el estrés hídrico (Cruzado-Vargas et al., 2019) y facilita el crecimiento (Valladares y Niinemets 2008). En particular, la ausencia de diferenciación en rasgos como el SLA, el largo de la acícula y la RGR muestra que, en los grupos (NR y RF), la captura de luz necesaria para producir biomasa para el crecimiento y elongación de órganos aéreos (Poorter 2002) respondió de forma similar a los pulsos de luz (Kitao et al., 2018; Mathur et al., 2018). Similar a lo aquí observado para *A. religiosa*, otras especies típicas de bosques mexicanos (i.e., *Pinus* y *Quercus*) de alta montaña son tolerantes a la sombra, y poseen adaptaciones que maximizan la fijación de luz y controlan la pérdida de agua (Sáenz-Romero et al., 2016; Martins et al., 2018).

La diferenciación fenotípica, probablemente ligada a la manipulación de los juveniles durante la plantación, abre la posibilidad de detectar el efecto de las presiones selectivas a nivel intrapoblacional (capítulo II). Al evaluar la influencia de los patrones genéticos y de la heterogeneidad del suelo en la arquitectura de los rasgos fenotípicos, se encontró que los primeros explicaron una mayor proporción de la varianza en todos los modelos (TVE de la Tabla S6). En el caso de la heterogeneidad del suelo, fueron ciertas variables específicas de fertilidad las que tuvieron influencia en la variación cuantitativa entre grupos de árboles (la disponibilidad de NH_4^+, TN, NO_3^-:NH_4^+ y TP). Es interesante notar que estos resultados también indican que los factores adaptativos no están afectando

de la misma forma a los grupos de individuos, a pesar de su escasa diferenciación genética (Fig. 1; capítulo IV). Aun así, dentro de los análisis, la mayor proporción de la variación fenotípica no pudo ser explicada (aproximadamente el 80% de la variación total). Por lo tanto, es probable que los rasgos medidos respondan a plasticidad fenotípica o a presiones selectivas que no se cuantificaron; dicho patrón ya se ha observado en otras especies forestales en condiciones naturales y controladas (Budde et al., 2014; Riordan et al., 2016; Talbot et al., 2017; de Villemereuil et al., 2018).

Los resultados obtenidos con nuestros modelos determinaron que al incluir las dos particiones de rasgos (crecimiento y fisiología), la predicción mejoró en comparación con los modelos de particiones individuales (Tabla 3; capítulo II). También se encontró que los dos factores edáficos con influencia más fuertes en la diferenciación de los grupos fueron las heterogeneidades de NH_4^+ y TN. Estos hallazgos, junto a las catorce asociaciones genotipo-fenotipo (Fig 5, capítulo II) y los valores de h_g^2 reportados (capítulo IV), sugieren que la selección natural estaría actuando sobre la evolución de la variación cuantitativa (σ_p^2) de los rasgos medidos; sobre todo teniendo en cuenta la anotación de dos de los genes candidatos, ambos asociados con la disponibilidad de nitrógeno microambiental (un gen de la super familia de las *pentatricopeptide repeat proteins* y el gen del *glycosyl hydrolase 9B13*; Opassiri et al., 2006; Spokevicius et al 2017).

Algunos estudios previos han mostrado una considerable respuesta genética a la variación en la composición nutricional y al estrés hídrico del suelo en especies forestales (p. ej., Eckert et al., 2012; Guerrero et al., 2018). En el capítulo IV se observaron diferencias de pH, conductividad eléctrica y nitrógeno total entre los sustratos de los dos sitios evaluados, siendo San Bartolo el sitio con condiciones ligeramente más fértiles para el crecimiento y desarrollo de las plantas de oyamel. Cabe anotar que no se analizó la correlación entre la variación fenotípica y microambiental en este lugar, y que aún hace falta determinar si las diferencias en diámetro de las plantas (Tabla 1), hasta qué punto se pueden explicar por la heterogeneidad del suelo entre sitios, y si se pueden encontrar variantes alélicas asociadas. La variación microgeográfica dentro (‘Rincón de Guadalupe’) y entre sitios (‘Rincón de Guadalupe’ vs. ‘San Bartolo’) es además un llamado a investigar de forma más generalizada los procesos selectivos del suelo a pequeñas escalas, y a buscar genes o regiones genómicas afectadas por dichos procesos si es posible en experimentos controlados.

Nuestros resultados apuntan también que existe un efecto de interacción entre el genotipo y el microambiente edáfico (G × E), que actúa sobre sobre una serie de rasgos fenotípicos de los individuos. Esto se confirma con el bajo ajuste predictivo en la validación cruzada (capítulo II) a partir de las variables del suelo. Resultados similares se obtuvieron con las aproximaciones bayesianas de predicción genómica (ver abajo). En otros árboles, predicciones de determinados rasgos fenotípicos en diversas condiciones ambientales también fueron bastante imprecisos, sugiriendo G × E (Resende et al., 2012; Cappa et al., 2016; Suontama et al., 2019). Por lo tanto, incluir la interacción microespacial en estudios de este tipo será fundamental, dadas las enormes presiones selectivas que puede ejercer el suelo sobre la supervivencia de los árboles (Eckert et al., 2012; Pregitzer et al., 2013). Sin embargo, muchos de los estudios en especies forestales se han desarrollado hasta el momento a escalas geográficas grandes (De Mita et al., 2013; Sork, 2018; Zhang et al., 2019), y sin considerar las presiones selectivas del suelo.

El proceso de reforestación de las plantas estudiadas ha aparentemente afectado de manera significativa el suelo del bosque, creando un mosaico de microambientes y presiones selectivas planta-suelo. Entre otras se sugiere una disponibilidad diferencial de TN y C (y en general de materia orgánica) entre los individuos NR y RF. Esto probablemente se explica por la alteración del suelo durante el trasplante (i.e., se retiró la capa superficial del mantillo y luego se mezclaron los horizontes superficiales del suelo) que pudo haber generado condiciones edáficas de crecimiento distintas en los individuos RF. Esto indica que las actividades humanas pueden afectar el balance químico natural del suelo, no solo antes del transplante (p. ej., mediante la inoculación de micorrizas y los tratamientos con fertilizantes; Le et al., 2012; Larkin et al., 2017; Bruelheide et al., 2018), sino también durante la plantación en campo. En estudios futuros, se sugiere cuantificar la estructura y función del microbioma asociada a la rizósfera, para investigar el efecto del mismo en el genotipo del individuo (G × E), así como la competencia entre grupos funcionales y otros factores.

Apoyándonos en la correlación entre la diferenciación genética y la variación fenotípica presentada en el capítulo II, junto a las características genético-poblacionales descritas en el capítulo III (p. ej., incremento en la consanguinidad, bajo N_e y baja diferenciación genética dentro de la muestra), se validó el uso de modelos de PG en un estudio piloto para probar la factibilidad de implementar estos modelos en plantas jóvenes de poblaciones naturales (capítulo IV). La idea es predecir el desempeño fenotípico de *A. religiosa* sin depender de información funcional *a priori* del efecto de los marcadores en

el rasgo de interés. Es decir, sin tener que buscar genes causales *a priori*, ya que las características genético-poblacionales arriba mencionadas favorecen la presencia de alelos compartidos entre individuos, que son la base teórica de los modelos de predicción genómica (Müller et al., 2017; Calleja-Rodríguez et al., 2020). En otras palabras, los patrones de diferenciación y parentesco indicarían que las poblaciones relictuales (como el caso de *A. religiosa*) pueden "imitar" las poblaciones fundadoras de los programas de mejoramiento, lo que a su vez permite crear modelos predictivos. Algunos ejemplos que podrían ser comparables a lo presentado en esta tesis son los programas de mejoramiento en especies con relaciones genéticas bajas (Resende et al., 2017; Hiraoka et al., 2018) y las plantaciones forestales durante su primer ciclo de mejoramiento (Rambolarimanana et al., 2018). Nuestros resultados indican que dichos modelos predictivos podrían aplicarse a poblaciones de otras especies con distribución fragmentada, como *Taxus baccata* en las regiones templadas de Eurasia o la gran mayoría de especies de alta montaña en México; incluso en poblaciones aisladas de especies más cosmopolitas, como las localidades *Pinus sylvestris* en el sur de España (Calleja-Rodríguez et al., 2020) o las de *Pseudotsuga menziesii* en el centro de México (Cruz-Nicolás et al., 2011). Todas ellas han mostrado una alta consanguinidad después de periodos prolongados de aislamiento geográfico (González-Martínez at al., 2010; Linares 2020). Aunque aún falta mucha información biológica para la mayoría de las especies forestales, nuestros resultados confirman enfoques prometedores para explorar rasgos adaptativos y que posiblemente acelerarán los programas reproductivos de los árboles (Eckert et al., 2013; Heslot et al., 2015; Lascoux et al., 2016).

En esta tesis también además se estudiaron modelos desarrollados para cada grupo de individuos, obteniéndose mayores heredabilidades y habilidades de predicción para los individuos regenerados naturalmente (NR y SB). En este sentido, es importante destacar que los modelos de PG se van a enfrentar a importantes desafíos que permitan la implementación en poblaciones naturales como son: el tamaño de muestra para el fenotipado y el genotipado del conjunto con el cual se entrena el modelo, así como la estructura genética subyacente (la cual debería ser similar entre TRN y TST, capítulo III; Isik et al., 2016; Grattapaglia et al., 2018), la edad de las población de entrenamiento (Ratcliffe et al., 2015), la presencia de los efectos de las interacciones G × E en respuesta a condiciones ambientales naturales en diferentes escalas espaciales (Resente et al., 2012; Ratcliffe et al., 2015), el tipo de métodos para desarrollar y probar el modelo (p. ej. GBLUP, RKHS, BRR, entre otros; Rodríguez-Pérez y de los Campos 2015), la cantidad

de marcadores que se incluyen en este (Thistletwaite et al., 2020) y la manera de evaluar su desempeño a largo plazo (Thistletwaite et al., 2019).

Las condiciones ambientales naturales de cada sitio parecen influenciar el desempeño de los individuos y sus efectos parecen ser predecibles a nivel local. Esto se sustenta en el modelo construido a partir de individuos regenerados naturalmente y utilizado en plantas introducidas en el mismo lugar (NR → RF), que fue el más eficiente de los modelos cruzados. Éste capturó mejor los efectos para ciertos rasgos (p. ej. G_{2016}, Ψ y SD; capítulo IV, Fig.3), indicando que los modelos de PG con diseño local deberían favorecerse en programas de translocación, para optimizar el rendimiento y supervivencia de los individuos a introducir. Por el contrario, los modelos construidos entre sitios fueron los menos eficientes (SB → NR y SB → RF; capítulo IV). Esto pudo deberse a que la interacción G × E afecta la precisión del modelo entre entornos, tal como ya se ha documentado en plantaciones forestales (p. ej. Resende et al. 2012, Lenz et al. 2017; Thistlewaite et al. 2019). Para ambientes tropicales de alta montaña, la interacción G × E microgeográfica estaría determinada por la variación en la cantidad de luz y los factores edáficos (ver capítulo II), que son dos recursos determinantes para el crecimiento, supervivencia y reproducción de las plantas (Pregitzer et al., 2013; Valladares et al., 2014; Li et al., 2018).

Apoyando la problemática de integrar los modelos de PG en las poblaciones naturales descrita en el capítulo III, se observa en nuestros resultados empíricos que incluir la interacción G × E es un desafío que complica la transferibilidad de modelos. Este mismo fenómeno también se ha reportado en otros estudios con especies de mejoramiento forestal (p. ej., Lenz et al., 2017; Ratcliffe et al., 2019). En nuestro caso, la habilidad de predicción de los modelos cruzados solamente mejoró cuando se incluyó germoplasma del entorno objetivo en el conjunto de entrenamiento (capítulo IV). Sin embargo, es necesario considerar que la diferenciación entre plantas solo se evidenció en individuos jóvenes, lo cual no implica que haya patrones de divergencia fenotípica a largo plazo. Lo anterior se podría solucionar, al menos parcialmente, con el diseño de experimentos de jardín común con diferentes procedencias y utilizando réplicas en distintas condiciones (Tiffin y Ross-Ibarra, 2014; Housset et al., 2018; Isabel et al., 2020). Esta aproximación podrá sin duda proporcionar una mejor comprensión de cómo la predictivilidad del fenotipo responde a la interacción G × E (Sork, 2018).

Los procesos adaptativos identificados en esta investigación pueden ser resultado de que las poblaciones naturales evaluadas aún mantienen las variantes alélicas

adaptativas necesarias para responder a las presiones selectivas ejercidas por el suelo (y probablemente de otros factores). Estas presiones son principalmente la composición nutricional y la asociación con comunidades del suelo. Una de las dificultades de medir adaptación local en coníferas tropicales como *A. religiosa* es cuantificar el *fitness* (adecuación biológica) a lo largo de su vida (Ortíz-Bibian et al., 2018). Para poder evaluarlo se necesitan estudios de largo plazo e incluir rasgos demográficos (Peterson et al., 2016). Una medida directa del *fitness* es la tasa de crecimiento poblacional, que integra la supervivencia, la reproducción, la capacidad para generar bancos de semillas y el crecimiento variable entre las etapas de la historia de vida (Mitchell -Olds et al., 2007; Blanquart et al., 2013; Valladares et al., 2014). En la presente tesis sólo se presentaron datos de crecimiento en plantas juveniles, por lo que aún haría falta incluir en las predicciones, rasgos de diferentes rangos de edades, particularmente de los adultos reproductivos, (Lascoux et al., 2016; Housset et al., 2018; Wachowiak et al., 2018). Esto representa un gran desafío para la mayoría de las investigaciones en genética forestal, que generalmente incluyen una sola generación (i.e., rango de edad).

Para comprender mejor la respuesta adaptativa de *A. religiosa* es necesario diseñar experimentos que permitan poner a prueba la hipótesis de adaptación local en rasgos directamente relacionados al *fitness*, e incluir diversas poblaciones junto a la variabilidad edáfica. La variabilidad del suelo detectada en dos localidades del Nevado de Toluca (Rincón de Guadalupe y San Bartolo) sugiere que los hábitats de *A. religiosa* son heterogéneos, por lo que habría que considerar un mayor número de variables que puedan contribuir en procesos adaptativos importantes, incluyendo el microclima, la composición de la vegetación, la competencia, y las interacciones bióticas, como las micorrizas del suelo (Schweitzer et al., 2011; Pregitzer et al., 2013; Li et al., 2018). Teniendo en cuenta tanto la ecología de los bosques de los *Abies* de México como la de los ecosistemas relictuales alpinos, algunas de las variables más valiosas podrían ser, por ejemplo, la exposición al viento y la cantidad de radiación solar incidente sobre la planta (i.e., efecto de ladera), ambas variables dependen de la ubicación microgeográfica dentro del bosque y se relacionan directamente con la cantidad de transpiración, y por lo tanto con el microclima.

Finalmente, sugerimos que el potencial predictivo de los modelos podría mejorar al utilizar estadística frecuentista y modelos de aprendizaje virtual (p. ej. GBLUP), y controlar el efecto G × E a escala microgeográfica (Kubota et al., 2015). Esto resalta la necesidad de cuantificar y caracterizar la variación ambiental a niveles microespaciales

en el momento de hacer reforestaciones y demás planes de manejo forestal, ya que los procesos ecológicos y/o evolutivos importantes para el buen desempeño de las plantas podrían estar ocurriendo a esta escala.

5.2. Perspectivas

A medida que las tecnologías de secuenciación de ADN continúen mejorando y evolucionando, aumentará la disponibilidad y la calidad de los recursos genómicos para organismos con genomas grandes y complejos, como las coníferas. Estas tecnologías permitirán explorar otros tipos de variación que afecten el fenotipo de las plantas, como la varianza epigenética y la variación en el número de copias de genes, que también podrían ser incluidas en modelos predictivos. Sin embargo, esto requerirá un estudio detallado de su heredabilidad (su naturaleza aditiva o no aditiva), ya que muchas de estas son únicamente fenotipos. Por ejemplo, la varianza epigenética puede usarse para mejorar rasgos valiosos y heredablemente estables, principalmente la tolerancia/resistencia al estrés (p. ej., hídrico o pestes). Se ha visto que cambios en la metilación de elementos transponibles pueden ser responsables de la variabilidad de algunos de estos rasgos (Sow et al., 2018; Wickler et al., 2018), que son especialmente importantes para especies que deberán enfrentar cambios ambientales futuros.

La adopción de técnicas combinadas de modelos poligénicos y de PG, será de vital importancia para evaluar los mecanismos y las bases genéticas de la adaptación a las condiciones del ambiente e identificar respuestas específicas ante eventos de estrés (capítulo III). Se prevé que la duración y la intensidad de factores estresantes va a ir aumentando durante los próximos años (p. ej. el calor y la sequía; Dawson et al., 2014), y a muchas poblaciones silvestres de árboles se les dificultará responder a tiempo a desafíos tan súbitos (Mitchell-Olds et al., 2007; Oddou-Muratorio et al., 2020; Santos y Gaiotto, 2020), lo que afectará su capacidad adaptativa (Lascoux et al., 2016; Sork, 2018). Usar estas técnicas sería facilitado por la expansión de herramientas genómicas en árboles forestales y por el número creciente de especies estudiadas; se espera que este número continúe aumentando en el futuro. Entre las especies boreales, algunos buenos candidatos a estudiar son especies de los géneros *Betula*, y de las tropicales en general. Las metodologías mencionadas pueden capturar relaciones entre rasgos y medio ambiente de múltiples escalas (Bruelheide et al., 2018) en especies arbóreas no modelo (Mayol et al.,

2020). Dada la complejidad y heterogeneidad de las fuentes de datos transdisciplinarias, *Machine Learning* (ML) y la conectividad entre estudios de asociación genómica con transcriptómica ofrecen enfoques predictivos y sintetizadores oportunos capaces de fusionar los aspectos más destacados de las técnicas ya nombradas (Depardiu et al., 2021). De esta manera, los modelos generados algorítmicamente ofrecerán nuevas formas de entender sistemas naturales muy complejos (p. ej. G x E; Myburg et al., 2019). De la misma manera en que ya se está utilizando en genómica funcional (Libbrecht y Noble, 2015) y modelado de nichos ecológicos (Phillips et al., 2017).

Para poder implementar modelos poligénicos y de predicción genómica para el manejo y conservación de poblaciones naturales de árboles, se necesitará un trabajo intensivo de fenotipado, que este es aún un proceso limitante. Se espera que en los siguientes años se amplíen las bases de datos de rasgos fisiológicos y de crecimiento en especies de árboles clave, como *A. religiosa*. La utilización de herramientas como la cromatografía líquida de alta resolución (*HPLC)*, de estudios de isótopos estables de carbono $\delta 1^3C$ y el intercambio de gases en el infrarrojo ayudarán a desarrollar modelos que incluyan rasgos sensibles al ambiente, como la resistencia a la sequía y otras presiones abióticas (Isabel et al., 2020; Lenz et al., 2020). Sin embargo, evaluar estos rasgos utilizando enfoques tradicionales es complicado y demorado, por lo que será necesario integrar tecnologías de fenotipificación masiva (p. ej. geomática y modelado basado en imágenes o fenotipificación de alto rendimiento; Fiorani and Schurr, 2013), que permitan una evaluación rápida, precisa y eficiente; por ejemplo, mediante la captura de imágenes en varias regiones espectrales (Lebedev et al., 2020). Las tecnologías de detección remota de alta resolución, son un ejemplo de ello; estas permiten obtener y procesar datos fenotípicos y ambientales a gran escala (Fiorani and Schurr, 2013), permitiendo monitorear fenotipos a través del tiempo.

Anticipamos que las técnicas de modelado con ML, CNV, transcriptómica y variación epigenética reforzarán las estimaciones de GP y de modelos poligénicos para varios rasgos en estudios multiambiente que tienen como objetivo desenredar la varianza genética aditiva y los componentes genotipo × entorno para entender la arquitectura genómica de la adaptación a las presiones del ambiente. Para esto, podemos esperar un uso más extenso de las simulaciones a nivel de implementación de modelos poligénicos y de PG. De esta manera, los nuevos desarrollos permitirán además construir predicciones más precisas mediante la fusión de variables ambientales, diversidad de microhábitats y

divergencia en todo el genoma, todo dentro de un contexto de adaptación al cambio climático en bosques naturales.

5.3 Referencias

Aguirre-Planter, E., Furnier, G. R., and Eguiarte, L. E. (2000). Low levels of genetic variation within and high levels of genetic differentiation among populations of species of Abies from southern Mexico and Guatemala. *Am. J. Bot.* 87, 362–371. doi:10.2307/2656632.

Aitken, S. N., and Bemmels, J. B. (2016). Time to get moving: Assisted gene flow of forest trees. *Evol. Appl.* 9, 271–290. doi:10.1111/eva.12293.

Aitken, S. N., and Whitlock, M. C. (2013). Assisted Gene Flow to Facilitate Local Adaptation to Climate Change. *Annu. Rev. Ecol. Evol. Syst.* 44, 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. *Evol. Appl.* 1, 95–111. doi:10.1111/j.1752-4571.2007.00013.x.

Alfaro, R. I., Fady, B., Vendramin, G. G., Dawson, I. K., Fleming, R. A., Sáenz-Romero, C., et al. (2014). The role of forest genetic resources in responding to biotic and abiotic factors in the context of anthropogenic climate change. *For. Ecol. Manage.* 333, 76–87. doi:10.1016/j.foreco.2014.04.006.

Anderson, J. T., Willis, J. H., and Mitchell-olds, T. (2012). NIH Public Access. 27, 258–266. doi:10.1016/j.tig.2011.04.001.Evolutionary.

Ángeles-Cervantes, Efraín López-Mata, L. (2009). SUPERVIVENCIA DE UNA COHORTE DE PLÁNTULAS DE ABIES RELIGIOSA BAJO DIFERENTES CONDICIONES POSTINCENDIO. *Bol.Soc.Bot.Méx* 84, 25–33.

Baker, P. J. (2003). Tree age estimation for the tropics: A test from the southern Appalachians. *Ecol. Appl.* 13, 1718–1732. doi:10.1890/02-5025.

Barrett, R. D. H., and Schluter, D. (2007). Adaptation from standing genetic variation. doi:10.1016/j.tree.2007.09.008.

Beaulieu, J., Nadeau, S., Ding, C., Celedon, J. M., Azaiez, A., Ritland, C., et al. (2020). Genomic selection for resistance to spruce budworm in white spruce and relationships with growth and wood quality traits. *Evol. Appl.*, 2704–2722. doi:10.1111/eva.13076.

Bezemer, N., Krauss, S. L., Roberts, D. G., and Hopper, S. D. (2019). Conservation of old individual trees and small populations is integral to maintain species' genetic diversity of a historically fragmented woody perennial. *Mol. Ecol.* 28, 3339–3357. doi:10.1111/mec.15164.

Biologicas, C., Juan, B., and Vicente, C. (2020). SAN NICOLAS DE HIDALGO Facultad de Biología.

Blanquart, F., Kaltz, O., Nuismer, S. L., and Gandon, S. (2013). A practical guide to measuring local adaptation. *Ecol. Lett.* 16, 1195–1205. doi:10.1111/ele.12150.

Bouvet, J. M., Makouanzi Ekomono, C. G., Brendel, O., Laclau, J. P., Bouillet, J. P., and Epron, D. (2020). Selecting for water use efficiency, wood chemical traits and biomass with genomic selection in a Eucalyptus breeding program. *For. Ecol. Manage.* 465, 118092. doi:10.1016/j.foreco.2020.118092.

Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S. M., et al. (2018). Global trait–environment relationships of plant communities. *Nat. Ecol. Evol.* 2, 1906–1917. doi:10.1038/s41559-018-0699-8.

Budde, K. B., Heuertz, M., Hernández-Serrano, A., Pausas, J. G., Vendramin, G. G., Verdú, M., et al. (2014). In situ genetic association for serotiny, a fire-related trait, in Mediterranean maritime pine (Pinus pinaster). *New Phytol.* 201, 230–241. doi:10.1111/nph.12483.

Calleja-Rodriguez, A., Pan, J., Funda, T., Chen, Z., Baison, J., Isik, F., et al. (2020). Evaluation of the efficiency of genomic versus pedigree predictions for growth and wood quality traits in Scots pine. *BMC Genomics* 21, 1–17. doi:10.1186/s12864-020-07188-4.

Capblancq, T., Luu, K., Blum, M. G. B., and Bazin, E. (2018). Evaluation of redundancy analysis to identify signatures of local adaptation. *Mol. Ecol. Resour.* 18, 1223–1233. doi:10.1111/1755-0998.12906.

Cappa, E. P., Stoehr, M. U., Xie, C. Y., and Yanchuk, A. D. (2016). Identification and joint modeling of competition effects and environmental heterogeneity in three Douglas-fir (Pseudotsuga menziesii var. menziesii) trials. *Tree Genet. Genomes* 12. doi:10.1007/s11295-016-1061-4.

Cappai, C., Kemanian, A. R., Lagomarsino, A., Roggero, P. P., Lai, R., Agnelli, A. E., et al. (2017). Small-scale spatial variation of soil organic matter pools generated by cork oak trees in

Mediterranean agro-silvo-pastoral systems. *Geoderma* 304, 59–67. doi:10.1016/j.geoderma.2016.07.021.
Carbajal-Navarro, A., Navarro-Miranda, E., Blanco-García, A., Cruzado-Vargas, A. L., Gómez-Pineda, E., Zamora-Sánchez, C., et al. (2019). Ecological Restoration of Abies religiosa Forests Using Nurse Plants and Assisted Migration in the Monarch Butterfly Biosphere Reserve, Mexico. *Front. Ecol. Evol.* 7. doi:10.3389/fevo.2019.00421.
Chen, J., Shen, W., Xu, H., Li, Y., and Luo, T. (2019). The composition of nitrogen-fixing microorganisms correlates with soil nitrogen content during reforestation: A comparison between legume and non-legume plantations. *Front. Microbiol.* 10, 1–11. doi:10.3389/fmicb.2019.00508.
Chen, Z. Q., Baison, J., Pan, J., Karlsson, B., Andersson, B., Westin, J., et al. (2018). Accuracy of genomic selection for growth and wood quality traits in two control-pollinated progeny trials using exome capture as the genotyping platform in Norway spruce. *BMC Genomics* 19. doi:10.1186/s12864-018-5256-y.
Cobo-Simón, I., Méndez-Cea, B., Jump, A. S., Seco, J., Gallego, F. J., and Linares, J. C. (2020). Understanding genetic diversity of relict forests. Linking long-term isolation legacies and current habitat fragmentation in Abies pinsapo Boiss. *For. Ecol. Manage.* 461, 117947. doi:10.1016/j.foreco.2020.117947.
Collevatti, R. G., Novaes, E., Silva-Junior, O. B., Vieira, L. D., Lima-Ribeiro, M. S., and Grattapaglia, D. (2019). A genome-wide scan shows evidence for local adaptation in a widespread keystone Neotropical forest tree. *Heredity (Edinb).* 123, 117–137. doi:10.1038/s41437-019-0188-0.
Cortés, A. J., Restrepo-Montoya, M., and Bedoya-Canas, L. E. (2020). Modern Strategies to Assess and Breed Forest Tree Adaptation to Changing Climate. *Front. Plant Sci.* 11. doi:10.3389/fpls.2020.583323.
Corvin, A., Craddock, N., and Sullivan, P. F. (2010). Genome-wide association studies: A primer. *Psychol. Med.* 40, 1063–1077. doi:10.1017/S0033291709991723.
Cruz-Nicolás, J., Giles-Pérez, G., González-Linares, E., Múgica-Gallart, J., Lira-Noriega, A., Gernandt, D. S., et al. (2019). Contrasting evolutionary processes drive morphological and genetic differentiation in a subtropical fir (Abies, Pinaceae) species complex. *Bot. J. Linn. Soc.*, 401–420. doi:10.1093/botlinnean/boz077.
Cruzado-Vargas, A. L., Zamudio-Sánchez, F. J., Rodríguez-Yam, G. A., Carbajal-Navarro, A. L., Blanco-García, J. A., and Sáenz-Romero, C. (2019). Growth of naturally regenerated Abies religiosa (Kunth) Schltdl. & Cham. Seedlings in a nursery and genetic variation among provenances. *Rev. Chapingo, Ser. Ciencias For. y del Ambient.* 26, 85–96. doi:10.5154/r.rchscfa.2019.01.013.
Dawson, I. K., Leakey, R., Clement, C. R., Weber, J. C., Cornelius, J. P., Roshetko, J. M., et al. (2014). The management of tree genetic resources and the livelihoods of rural communities in the tropics: Non-timber forest products, smallholder agroforestry practices and tree commodity crops. *For. Ecol. Manage.* 333, 9–21. doi:10.1016/j.foreco.2014.01.021.
De Mita, S., Thuillet, A. C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., et al. (2013). Detecting selection along environmental gradients: Analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Mol. Ecol.* 22, 1383–1399. doi:10.1111/mec.12182.
de Villemereuil, P., Mouterde, M., Gaggiotti, O. E., and Till-Bottraud, I. (2018). Patterns of phenotypic plasticity and local adaptation in the wide elevation range of the alpine plant Arabis alpina. *J. Ecol.* 106, 1952–1971. doi:10.1111/1365-2745.12955.
Eckert, A. J., Shahi, H., Datwyler, S. L., and Neale, D. B. (2012). Spatially variable natural selection and the divergence between parapatric subspecies of lodgepole pine (Pinus contorta, Pinaceae). *Am. J. Bot.* 99, 1323–1334. doi:10.3732/ajb.1200055.
Eckert, A. J., Wegrzyn, J. L., Liechty, J. D., Lee, J. M., Cumbie, W. P., Davis, J. M., et al. (2013). The evolutionary genetics of the genes underlying phenotypic associations for loblolly pine (Pinus taeda, Pinaceae). *Genetics* 195, 1353–1372. doi:10.1534/genetics.113.157198.
Elliot, W. J., Page-Dumroese, D., and Robichaud, P. R. (2019). The Effects of Forest Management on Erosion and Soil Productivity*. *Soil Qual. Soil Eros.*, 195–208. doi:10.1201/9780203739266-12.
Ellison, D., Morris, C. E., Locatelli, B., Sheil, D., Cohen, J., Murdiyarso, D., et al. (2017). Trees, forests and water: Cool insights for a hot world. *Glob. Environ. Chang.* 43, 51–61. doi:10.1016/j.gloenvcha.2017.01.002.
Finlay, C. M. V., Bradley, C. R., Jane Preston, S., and Provan, J. (2017). Low genetic diversity and potential inbreeding in an isolated population of alder buckthorn (Frangula alnus) following a founder effect. *Sci. Rep.* 7, 1–8. doi:10.1038/s41598-017-03166-1.
Fiorani, F., and Schurr, U. (2013). Future scenarios for plant phenotyping. *Annu. Rev. Plant Biol.* 64, 267–291. doi:10.1146/annurev-arplant-050312-120137.
George, J. P., Grabner, M., Karanitsch-Ackerl, S., Mayer, K., Weißenbacher, L., and Schueler, S. (2017).

Genetic variation, phenotypic stability, and repeatability of drought response in European larch throughout 50 years in a common garden experiment. *Tree Physiol.* 37, 33–46. doi:10.1093/treephys/tpw085.
Gernandt, D. S., and Pérez-De La Rosa, J. A. (2014). Biodiversidad de Pinophyta (coníferas) en México. *Rev. Mex. Biodivers.* 85, 126–133. doi:10.7550/rmb.32195.
Gömöry, D., Comps, B., Paule, L., and Von Wühlisch, G. (2013). Allozyme and phenotypic variation in beech (Fagus sylvatica L.): Are there any links? *Plant Biosyst.* 147, 265–271. doi:10.1080/11263504.2013.763864.
Grattapaglia, D. (2014). *Genomics of plant genetic resources: Volume 1. Managing, sequencing and mining genetic resources. Chapter 26 Breeding Forest Trees by Genomic Selection: Current Progress and the Way Forward*. doi:10.1007/978-94-007-7572-5.
Guerrero, J., Andrello, M., Burgarella, C., and Manel, S. (2018). Soil environment is a key driver of adaptation in Medicago truncatula: new insights from landscape genomics. *New Phytol.* 219, 378–390. doi:10.1111/nph.15171.
Habier, D., Tetens, J., Seefried, F. R., Lichtner, P., and Thaller, G. (2010). The impact of genetic relationship information on genomic breeding values in German Holstein cattle. *Genet. Sel. Evol.* 42, 1–12. doi:10.1186/1297-9686-42-5.
Harter, D. E. V., Nagy, L., Backhaus, S., Beierkuhnlein, C., Fussi, B., Huber, G., et al. (2015). A comparison of genetic diversity and phenotypic plasticity among european beech (Fagus sylvatica L.) populations from bulgaria and germany under drought and temperature manipulation. *Int. J. Plant Sci.* 176, 232–244. doi:10.1086/679349.
Hatfield, J. L., and Dold, C. (2019). Water-use efficiency: Advances and challenges in a changing climate. *Front. Plant Sci.* 10, 1–14. doi:10.3389/fpls.2019.00103.
Hay, E. H. A., Utsunomiya, Y. T., Xu, L., Zhou, Y., Neves, H. H. R., Carvalheiro, R., et al. (2018). Genomic predictions combining SNP markers and copy number variations in Nellore cattle. *BMC Genomics* 19, 1–8. doi:10.1186/s12864-018-4787-6.
Hayes, B. J., Bowman, P. J., Chamberlain, A. J., and Goddard, M. E. (2009). Invited review : Genomic selection in dairy cattle : Progress and challenges. *J. Dairy Sci.* 92, 433–443. doi:10.3168/jds.2008-1646.
Heslot, N., Jannink, J.-L., and Sorrells, M. E. (2015). Perspectives for Genomic Selection Applications and Research in Plants. *Crop Sci.* 55, 1. doi:10.2135/cropsci2014.03.0249.
Hiraoka, Y., Fukatsu, E., Mishima, K., Hirao, T., Teshima, K. M., Tamura, M., et al. (2018). Potential of genome-wide studies in unrelated plus trees of a coniferous species, cryptomeria japonica (japanese cedar). *Front. Plant Sci.* 9, 1–15. doi:10.3389/fpls.2018.01322.
Holliday, J. A., Zhou, L., Bawa, R., Zhang, M., and Oubida, R. W. (2016). Evidence for extensive parallelism but divergent genomic architecture of adaptation along altitudinal and latitudinal gradients in Populus trichocarpa. *New Phytol.* 209, 1240–1251. doi:10.1111/nph.13643.
Housset, J. M., Nadeau, S., Isabel, N., Depardieu, C., Duchesne, I., Lenz, P., et al. (2018). Tree rings provide a new class of phenotypes for genetic associations that foster insights into adaptation of conifers to climate change. *New Phytol.* 218, 630–645. doi:10.1111/nph.14968.
Isabel, N., Holliday, J. A., and Aitken, S. N. (2020). Forest genomics: Advancing climate adaptation, forest health, productivity, and conservation. *Evol. Appl.* 13, 3–10. doi:10.1111/eva.12902.
Isik, F., Bartholomé, J., Farjat, A., Chancerel, E., Raffin, A., Sanchez, L., et al. (2016). Plant Science Genomic selection in maritime pine. *Plant Sci.* 242, 108–119. doi:10.1016/j.plantsci.2015.08.006.
Isik, F., Kumar, S., Martínez-García, P. J., Iwata, H., and Yamamoto, T. (2015). *Acceleration of Forest and Fruit Tree Domestication by Genomic Selection*. doi:10.1016/bs.abr.2015.05.002.
Kim, J. D., Lim, J. H., and Yun, C. W. (2019). Dynamics of Abies nephrolepis seedlings in relation to environmental factors in Seorak mountain, south Korea. *Forests* 10, 1–14. doi:10.3390/f10080702.
Kitao, M., Kitaoka, S., Harayama, H., Tobita, H., Agathokleous, E., and Utsugi, H. (2018). Canopy nitrogen distribution is optimized to prevent photoinhibition throughout the canopy during sun flecks. *Sci. Rep.* 8, 1–11. doi:10.1038/s41598-017-18766-0.
Koskela, J., Vinceti, B., Dvorak, W., Bush, D., Dawson, I. K., Loo, J., et al. (2014). Utilization and transfer of forest genetic resources: A global review. *For. Ecol. Manage.* 333, 22–34. doi:10.1016/j.foreco.2014.07.017.
Kremer, A., Potts, B. M., and Delzon, S. (2014). Genetic divergence in forest trees: Understanding the consequences of climate change. *Funct. Ecol.* 28, 22–36. doi:10.1111/1365-2435.12169.
Kubota, S., Iwasaki, T., Hanada, K., Nagano, A. J., Fujiyama, A., Toyoda, A., et al. (2015). A Genome Scan for Genes Underlying Microgeographic-Scale Local Adaptation in a Wild Arabidopsis Species. *PLoS Genet.* 11, 1–26. doi:10.1371/journal.pgen.1005361.
Larkin, R. P., Honeycutt, C. W., Griffin, T. S., Olanya, O. M., He, Z., and Halloran, J. M. (2017).

Cumulative and residual effects of different potato cropping system management strategies on soilborne diseases and soil microbial communities over time. *Plant Pathol.* 66, 437–449. doi:10.1111/ppa.12584.
Lascoux, M., Glémin, S., and Savolainen, O. (2016). Local Adaptation in Plants. *eLS*, 1–7. doi:10.1002/9780470015902.a0025270.
Lawson, S. S., and Michler, C. H. (2014). Afforestation, restoration and regeneration - Not all trees are created equal. *J. For. Res.* 25, 3–20. doi:10.1007/s11676-014-0426-5.
Le, H. D., Smith, C., Herbohn, J., and Harrison, S. (2012). More than just trees: Assessing reforestation success in tropical developing countries. *J. Rural Stud.* 28, 5–19. doi:10.1016/j.jrurstud.2011.07.006.
Lebedev, V. G., Lebedeva, T. N., Chernodubov, A. I., and Shestibratov, K. A. (2020). Genomic selection for forest tree improvement: Methods, achievements and perspectives. *Forests* 11, 1–36. doi:10.3390/f11111190.
Leimu, R., and Fischer, M. (2008). A meta-analysis of local adaptation in plants. *PLoS One* 3, 1–8. doi:10.1371/journal.pone.0004010.
Lenz, P. R. N., Beaulieu, J., Mansfield, S. D., Clément, S., Desponts, M., and Bousquet, J. (2017). Factors affecting the accuracy of genomic selection for growth and wood quality traits in an advanced-breeding population of black spruce (Picea mariana). *BMC Genomics* 18, 1–17. doi:10.1186/s12864-017-3715-5.
Lenz, P. R. N., Nadeau, S., Mottet, M. J., Perron, M., Isabel, N., Beaulieu, J., et al. (2020). Multi-trait genomic selection for weevil resistance, growth, and wood quality in Norway spruce. *Evol. Appl.* 13, 76–94. doi:10.1111/eva.12823.
Li, Y., Wu, X., Chen, T., Wang, W., Liu, G., Zhang, W., et al. (2018). Plant phenotypic traits eventually shape its microbiota: A common garden test. *Front. Microbiol.* 9, 1–13. doi:10.3389/fmicb.2018.02479.
Li Y, Dungey HS, Carson M, Carson S. Genotype by environment interaction for growth and Dothistroma resistance and clonal connectivity between environments in radiata pine in New Zealand and Australia. PLoS One. 2018;13(10):e0205402. Published 2018 Oct 12. doi:10.1371/journal.pone.0205402
Lin, Y. Te, Whitman, W. B., Coleman, D. C., and Chiu, C. Y. (2018). Effects of reforestation on the structure and diversity of bacterial communities in subtropical low Mountain Forest Soils. *Front. Microbiol.* 9, 1–10. doi:10.3389/fmicb.2018.01968.
Liu, M., Sui, X., Hu, Y., and Feng, F. (2019). Microbial community structure and the relationship with soil carbon and nitrogen in an original Korean pine forest of Changbai Mountain, China. *BMC Microbiol.* 19, 218. doi:10.1186/s12866-019-1584-6.
Lojewski, N. R., Fischer, D. G., Bailey, J. K., Schweitzer, J. A., Whitham, T. G., and Hart, S. C. (2009). Genetic basis of aboveground productivity in two native Populus species and their hybrids. *Tree Physiol.* 29, 1133–1142. doi:10.1093/treephys/tpp046.
Lotterhos, K. E., and Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol. Ecol.* 24, 1031–1046. doi:10.1111/mec.13100.
Lu, X., Mo, J., Gilliam, F. S., Fang, H., Zhu, F., Fang, Y., et al. (2012). Nitrogen Addition Shapes Soil Phosphorus Availability in Two Reforested Tropical Forests in Southern China. *Biotropica* 44, 302–311. doi:10.1111/j.1744-7429.2011.00831.x.
Madritch, M. D., and Lindroth, R. L. (2011). Soil microbial communities adapt to genetic variation in leaf litter inputs. *Oikos* 120, 1696–1704. doi:10.1111/j.1600-0706.2011.19195.x.
Martins, K., Gugger, P. F., Llanderal-Mendoza, J., González-Rodríguez, A., Fitz-Gibbon, S. T., Zhao, J. L., et al. (2018). Landscape genomics provides evidence of climate-associated genetic variation in Mexican populations of Quercus rugosa. *Evol. Appl.* 11, 1842–1858. doi:10.1111/eva.12684.
Mathur, S., Jain, L., and Jajoo, A. (2018). Photosynthetic efficiency in sun and shade plants. *Photosynthetica* 56, 354–365. doi:10.1007/s11099-018-0767-y.
Meuwissen, T., Hayes, B., and Goddard, M. (2016). Genomic selection: A paradigm shift in animal breeding. *Anim. Front.* 6, 6–14. doi:10.2527/af.2016-0002.
Mitchell-Olds, T., Willis, J. H., and Goldstein, D. B. (2007). Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nat. Rev. Genet.* 8, 845–856. doi:10.1038/nrg2207.
Montaño, N. & Sánchez, J (2014). Terra (1). 23, 98–104.
Müller, B. S. F., Neves, L. G., de Almeida Filho, J. E., Resende, M. F. R., Muñoz, P. R., dos Santos, P. E. T., et al. (2017). Genomic prediction in contrast to a genome-wide association study in explaining heritable variation of complex growth traits in breeding populations of Eucalyptus. *BMC Genomics* 18, 1–17. doi:10.1186/s12864-017-3920-2.

Negassa, W., Baum, C., Schlichting, A., Müller, J., and Leinweber, P. (2019). Small-Scale Spatial Variability of Soil Chemical and Biochemical Properties in a Rewetted Degraded Peatland. *Front. Environ. Sci.* 7, 1–15. doi:10.3389/fenvs.2019.00116.

Oddou-Muratorio, S., Davi, H., and Lefèvre, F. (2020). Integrating evolutionary, demographic and ecophysiological processes to predict the adaptive dynamics of forest tree populations under global change. *Tree Genet. Genomes* 16, 1–22. doi:10.1007/s11295-020-01451-1.

Osuna, F., González, D., de los Monteros, A. E., and Guerrero, J. A. (2020). Phylogeography of the Volcano Rabbit (Romerolagus diazi): the Evolutionary History of a Mountain Specialist Molded by the Climatic-Volcanism Interaction in the Central Mexican Highlands. *J. Mamm. Evol.* 27, 745–757. doi:10.1007/s10914-019-09493-6.

Paluch, J. G., and Jastrzebski, R. (2013). Natural regeneration of shade-tolerant Abies alba Mill. in gradients of stand species compositions: Limitation by seed availability or safe microsites? *For. Ecol. Manage.* 307, 322–332. doi:10.1016/j.foreco.2013.06.035.

Pautasso, M. (2009). Geographical genetics and the conservation of forest trees. *Perspect. Plant Ecol. Evol. Syst.* 11, 157–189. doi:10.1016/j.ppees.2009.01.003.

Peterson, M. L., Kay, K. M., and Angert, A. L. (2016). The scale of local adaptation in Mimulus guttatus: Comparing life history races, ecotypes, and populations. *New Phytol.* 211, 345–356. doi:10.1111/nph.13971.

Plomion, C., Bartholomé, J., Bouffier, L., Brendel, O., Cochard, H., De Miguel, M., et al. (2016). Understanding the genetic bases of adaptation to soil water deficit in trees through the examination of water use efficiency and cavitation resistance: maritime pine as a case study. *J. Plant Hydraul.* 3, 008. doi:10.20870/jph.2016.e008.

Poorter, H. (2002). Plant Growth and Carbon Economy. *Encycl. Life Sci.* doi:10.1038/npg.els.0003200.

Poorter, L., and Bongers, F. (2006). Leaf traits are good predictors of plant performance across 53 rain forest species. *Ecology* 87, 1733–1743. doi:10.1890/0012-9658(2006)87[1733:LTAGPO]2.0.CO;2.

Pregitzer, C. C., Bailey, J. K., and Schweitzer, J. A. (2013). Genetic by environment interactions affect plant-soil linkages. *Ecol. Evol.* 3, 2322–2333. doi:10.1002/ece3.618.

Pugnaire, F. I., Morillo, J. A., Peñuelas, J., Reich, P. B., Bardgett, R. D., Gaxiola, A., et al. (2019). Climate change effects on plant-soil feedbacks and consequences for biodiversity and functioning of terrestrial ecosystems. *Sci. Adv.* 5, eaaz1834. doi:10.1126/sciadv.aaz1834.

Qiu, W., Curtin, D., Johnstone, P., Beare, M., and Hernandez-Ramirez, G. (2016). Small-Scale Spatial Variability of Plant Nutrients and Soil Organic Matter: An Arable Cropping Case Study. *Commun. Soil Sci. Plant Anal.* 47, 2189–2199. doi:10.1080/00103624.2016.1228945.

Rambolarimanana, T., Ramamonjisoa, L., Verhaegen, D., Leong Pock Tsy, J. M., Jacquin, L., Cao-Hamadou, T. V., et al. (2018). Performance of multi-trait genomic selection for Eucalyptus robusta breeding program. *Tree Genet. Genomes* 14. doi:10.1007/s11295-018-1286-5.

Ratcliffe, B., Thistlethwaite, F. R., El-Dien, O. G., Cappa, E., Porth, I., Klápště, J., et al. (2019). Inter- and Intra-Generation Genomic Predictions for Douglas-fir Growth in Unobserved Environments. *bioRxiv*, 540765. doi:10.1101/540765.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., and Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Mol. Ecol.* 24, 4348–4370. doi:10.1111/mec.13322.

Resende, M. D. V., Resende, M. F. R., Sansaloni, C. P., Petroli, C. D., Missiaggia, A. A., Aguiar, A. M., et al. (2012a). Genomic selection for growth and wood quality in Eucalyptus: Capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytol.* 194, 116–128. doi:10.1111/j.1469-8137.2011.04038.x.

Resende, M. F. R., Muñoz, P., Acosta, J. J., Peter, G. F., Davis, J. M., Grattapaglia, D., et al. (2012b). Accelerating the domestication of trees using genomic selection: Accuracy of prediction models across ages and environments. *New Phytol.* 193, 617–624. doi:10.1111/j.1469-8137.2011.03895.x.

Resende, R. T., Resende, M. D. V, Silva, F. F., Azevedo, C. F., Takahashi, E. K., Silva-junior, O. B., et al. (2017). Assessing the expected response to genomic selection of individuals and families in Eucalyptus breeding with an additive-dominant model. 119, 245–255. doi:10.1038/hdy.2017.37.

Richardson, D. M., Hellmann, J. J., Mclachlan, J. S., Sax, D. F., Schwartz, M. W., Gonzalez, P., et al. (2009). Richardson et al. (2009)_Multidimensional evaluation of managed relocation.pdf. 1–4. doi:10.1073/pnas.0902327106.

Riordan, E. C., Gugger, P. F., Ortego, J., Smith, C., Gaddis, K., Thompson, P., et al. (2016). Association of genetic and phenotypic variability with geography and climate in three southern California oaks. *Am. J. Bot.* 103, 73–85. doi:10.3732/ajb.1500135.

Sáenz-Romero, C., Lindig-Cisneros, R. A., Joyce, D. G., Beaulieu, J., Bradley, J. S. C., and Jaquish, B. C. (2016). Assisted migration of forest populations for adapting trees to climate change. *Rev.*

Chapingo, Ser. Ciencias For. y del Ambient. 22, 303–323. doi:10.5154/r.rchscfa.2014.10.052.
Sasaki, S. (2008). Physiological characteristics of tropical rain forest tree species: A basis for the development of silvicultural technology. *Proc. Japan Acad. Ser. B Phys. Biol. Sci.* 84, 31–57. doi:10.2183/pjab.84.31.
Sauer, T. J., James, D. E., Cambardella, C. A., and Hernandez-Ramirez, G. (2012). Soil properties following reforestation or afforestation of marginal cropland. *Plant Soil* 360, 375–390. doi:10.1007/s11104-012-1258-8.
Savolainen, O., Pyhäjärvi, T., and Knürr, T. (2007). Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proc. Natl. Acad. Sci. U. S. A.* 38, 1752–1756. doi:10.1073/pnas.0906182107.
Schimel, J. P., and Bennett, J. (2004). Nitrogen mineralization: Challenges of a changing paradigm. *Ecology* 85, 591–602. doi:10.1890/03-8002.
Schueler, S., Kapeller, S., Konrad, H., Geburek, T., Mengl, M., Bozzano, M., et al. (2013). Adaptive genetic diversity of trees for forest conservation in a future climate: A case study on Norway spruce in Austria. *Biodivers. Conserv.* 22, 1151–1166. doi:10.1007/s10531-012-0313-3.
Schweitzer, J. A., Fischer, D. G., Rehill, B. J., Wooley, S. C., Woolbright, S. A., Lindroth, R. L., et al. (2011). Forest gene diversity is correlated with the composition and function of soil microbial communities. *Popul. Ecol.* 53, 35–46. doi:10.1007/s10144-010-0252-3.
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? *Ann. For. Sci.* 73, 69–75. doi:10.1007/s13595-015-0471-z.
Senneville, S., Beaulieu, J., Daoust, G., Deslauriers, M., and Bousquet, J. (2001). Evidence for low genetic diversity and metapopulation structure in Canada yew (Taxus canadensis): Considerations for conservation. *Can. J. For. Res.* 31, 110–116. doi:10.1139/cjfr-31-1-110.
Shen, Y., Wang, L., Fu, J., Xu, X., Yue, G. H., and Li, J. (2019). Population structure, demographic history and local adaptation of the grass carp. *BMC Genomics* 20, 1–16. doi:10.1186/s12864-019-5872-1.
Siefert, A., Violle, C., Chalmandrier, L., Albert, C. H., Taudiere, A., Fajardo, A., et al. (2015). A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecol. Lett.* 18, 1406–1419. doi:10.1111/ele.12508.
Sork, V. L. (2018). Genomic Studies of Local Adaptation in Natural Plant Populations. *J. Hered.* 109, 3–15. doi:10.1093/jhered/esx091.
Sow, M. D., Allona, I., Ambroise, C., Conde, D., Fichot, R., Gribkova, S., et al. (2018). Epigenetics in Forest Trees: State of the Art and Potential Implications for Breeding and Management in a Context of Climate Change. *Adv. Bot. Res.* 88, 387–453. doi:10.1016/bs.abr.2018.09.003.
Stocks, J. J., Metheringham, C. L., Plumb, W., Lee, S. J., Kelly, L. J., Nichols, R. A., et al. (2019). *Genomic basis of European ash tree resistance to ash dieback fungus*. doi:10.1101/626234.
Suontama, M., Mckinley, R., Dungey, H., Telfer, E., Graham, N., Stovold, T., et al. (2019). Ef fi ciency of genomic prediction across two Eucalyptus nitens seed orchards with different selection histories. 370–379. doi:10.1038/s41437-018-0119-5.
Talbot, B., Chen, T. W., Zimmerman, S., Joost, S., Eckert, A. J., Crow, T. M., et al. (2017). Combining genotype, phenotype, and environment to infer potential candidate genes. *J. Hered.* 108, 207–216. doi:10.1093/jhered/esw077.
Thistlethwaite, F. R., Ratcliffe, B., Klápště, J., Porth, I., Chen, C., Stoehr, M. U., et al. (2019). Genomic selection of juvenile height across a single-generational gap in Douglas-fir. *Heredity (Edinb).* 122, 848–863. doi:10.1038/s41437-018-0172-0.
Tiffin, P., and Ross-Ibarra, J. (2014). Advances and limits of using population genetics to understand local adaptation. *Trends Ecol. Evol.* 29, 673–680. doi:10.1016/j.tree.2014.10.004.
Valladares, F., Matesanz, S., Guilhaumon, F., Araújo, M. B., Balaguer, L., Benito-Garzón, M., et al. (2014). The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecol. Lett.* 17, 1351–1364. doi:10.1111/ele.12348.
Valladares, F., and Niinemets, Ü. (2008). Shade tolerance, a key plant feature of complex nature and consequences. *Annu. Rev. Ecol. Evol. Syst.* 39, 237–257. doi:10.1146/annurev.ecolsys.39.110707.173506.
Van Eenennaam, A. L., Weigel, K. A., Young, A. E., Cleveland, M. A., and Dekkers, J. C. M. (2014). Applied Animal Genomics : Results from the Field. *Annu. Rev. Anim. Biosci.* 2, 105–139. doi:10.1146/annurev-animal-022513-114119.
Wachowiak, W., Perry, A., Donnelly, K., and Cavers, S. (2018). Early phenology and growth trait variation in closely related European pine species. *Ecol. Evol.* 8, 655–666. doi:10.1002/ece3.3690.
Wang, J., Lu, N., Yi, F., and Xiao, Y. (2020). Identification of Transposable Elements in Conifer and

Their Potential Application in Breeding. *Evol. Bioinforma.* 16. doi:10.1177/1176934320930263.
Way, D. A., and Pearcy, R. W. (2012). Sunflecks in trees and forests: From photosynthetic physiology to global change biology. *Tree Physiol.* 32, 1066–1081. doi:10.1093/treephys/tps064.
Wicker, T., Gundlach, H., Spannagl, M., Uauy, C., Borrill, P., Ramírez-González, R. H., et al. (2018). Impact of transposable elements on genome structure and evolution in bread wheat. *bioRxiv*, 1–18. doi:10.1101/363192.
Zapata-Valenzuela, J., Whetten, R. W., Neale, D., McKeand, S., and Isik, F. (2013). Genomic Estimated Breeding Values Using Genomic Relationship Matrices in a Cloned Population of Loblolly Pine. *G3: Genes|Genomes|Genetics* 3, 909–916. doi:10.1534/g3.113.005975.
Zhang, M., Suren, H., and Holliday, J. A. (2019). Phenotypic and Genomic Local Adaptation across Latitude and Altitude in Populus trichocarpa. *Genome Biol. Evol.* 11, 2256–2272. doi:10.1093/gbe/evz151.
Zhou, L., Saeed, S., Sun, Y., Zhang, B., Luo, M., Li, Z., et al. (2019). The relationships between water storage and biomass components in two conifer species. *PeerJ* 2019. doi:10.7717/peerj.7901.