Universidad de Valladolid

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍAS AGRARIAS

INSTITUTO UNIVERSITARIO DE INVESTIGACIÓN
EN GESTIÓN FORESTAL SOSTENIBLE

Integrating infra-specific variation of Mediterranean conifers in species distribution models. Applications for vulnerability assessment and conservation

(Integración de la variación infra-específica de coníferas Mediterráneas en modelos de distribución de especies. Aplicaciones para la evaluación de la vulnerabilidad y la conservación)

Presentada por **Mª Jesús Serra Varela**
para optar al grado de Doctora por la
Universidad de Valladolid

Dirigida por el **Doctor Julián Gonzalo Jiménez**
y por la **Doctora Delphine Grivet**

GOBIERNO
DE ESPAÑA

MINISTERIO
DE CIENCIA
E INNOVACIÓN

INIA
Instituto Nacional de Investigación
y Tecnología Agraria y Alimentaria

Universidad de Valladolid

Instituto
Universitario de Investigación
iuFOR GESTIÓN
FORESTAL
SOSTENIBLE

Universidad de Valladolid

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍAS AGRARIAS

INSTITUTO UNIVERSITARIO DE INVESTIGACIÓN
EN GESTIÓN FORESTAL SOSTENIBLE

TESIS DOCTORAL:

Integrating infra-specific variation of Mediterranean conifers in species distribution models. Applications for vulnerability assessment and conservation

(Integración de la variación infra-específica de coníferas Mediterráneas en modelos de distribución de especies. Aplicaciones para la evaluación de la vulnerabilidad y la conservación)

Presentada por Mª Jesús Serra Varela
para optar al grado de Doctora por la Universidad de Valladolid

Dirigida por el Doctor Julián Gonzalo Jiménez
y por la Doctora Delphine Grivet

All models are wrong but some are useful.

George Box.

Agradecimientos

La tesis doctoral convive contigo un largo periodo de tiempo (entre 4-5 años en mi caso) abarcando muchos momentos de la vida distintos. Durante este tiempo, un gran número de personas se cruzan en tu camino y se involucran de una manera u otra dejando un poso, que poco a poco nos va definiendo y va marcando nuestra trayectoria. Es a todas estas personas a las que me gustaría agradecer que hayan hecho que mi camino haya sido así de increíble y enriquecedor. He tenido mucha suerte con todos aquellos que me habéis acompañado ¡MUCHAS GRACIAS!

Agradezco al Ministerio de Educación, Cultura y Deporte, la oportunidad que he me ha dado para crecer profesionalmente gracias a una beca para la Formación de Profesorado Universitario (FPU). Este agradecimiento es extensivo tanto a la Universidad de Valladolid, mi primer centro de adscripción, el centro donde hice el master, y donde he dado mis primeros pasos como profesora universitaria como al INIA, por permitirme trabajar en sus instalaciones desde el primer momento, y por acogerme durante la segunda fase de contrato. Además, quisiera agradecer su papel de vínculo entre ambas instituciones al Instituto Universitario para la Gestión Forestal Sostenible.

A mi director Julián Gonzalo, me gustaría agradecerle el haber confiado en mí sin conocerme, y haberme introducido en el mundo de la investigación. Primero con la oportunidad de trabajar en el Valle del Nansa, y después animándome y apoyándome hasta conseguir la beca de doctorado. También la libertad y autonomía que he tenido desde el principio y durante todo el desarrollo del trabajo. A mi directora, Delphine Grivet, quiero darle las gracias por su seguimiento durante toda la tesis, por sacar lo mejor de mí, por enseñarme a ser investigadora y a escribir, por leerse cada coma de mi trabajo y cuidar de todos los detalles con una sonrisa y con palabras de ánimo. Así, hacer la tesis ha sido una maravilla. Además, a Ricardo Alía, mi tutor en el INIA, gracias por todo el tiempo invertido en mejorar mi trabajo y en explicarme conceptos y por ser para mí un ejemplo de vocación investigadora.

I would also like to thank the COST ACTION FP1202 for the funding that enabled me to develop two Short Term Scientific Missions in Zürich. Thanks are extended to the Swiss Federal Institute for Forest, Snow and Landscape Research WSL for welcoming me during these stays. In particular, I would like to thank Niklaus E.

Zimmermann, as he guided my first steps using species distribution models to answer ecological questions, and as he maintained very fruitful discussions with me at any time - specially early in the morning ;) From the WSL, I would also like to thank Achilleas Psomas for his help and advice with R and GIS, and for being another Mediterranean person surviving among Swiss people ;) And of course, Bianca Saladin, Laure Gallien, Damaris Zurell, Stephan and Arnould for making life much nicer and funnier in the WSL, in Zürich and in the IBS Conference. From Zürich, I will not forget Tina, that welcomed me as a friend from the very beginning, showed me great places in Zürich, and even welcomed me in her house when necessary.

Por otra parte, me gustaría agradecer a GBIF el curso de modelos de nicho ecológico que imparte todos los años. En particular y muy especialmente a Blas Benito, por hacer un curso tan bueno, por sus ganas de enseñar, por poner a disposición de todos los participantes sus scripts, y sobre todo, por permitir un seguimiento del curso a distancia y en directo. Para mí, este curso fue decisivo y marcó un antes y un después en el desarrollo de la tesis.

Como no agradecer a Celia Redondo Marina, de la Unidad de Apoyo a Departamentos de la ETSIAA de la Universidad de Valladolid (Campus de Palencia), toda su ayuda con los trámites administrativos, por su paciencia, por estar siempre enterada de todo y por ayudarme a organizarme tan bien.

En estos momentos, me acuerdo también de Alfredo Blanco de la ETSI de Montes de Madrid, a quien quiero agradecerle su apoyo y su confianza en mi trabajo. Hoy es una de las personas, a las que me siento orgullosa de decir que voy a depositar la tesis doctoral.

También, quiero agradecerles a Esther y David, los profes de Juan, su vocación y entrega al desarrollar su trabajo. Me ha permitido centrarme 100 % durante mis horas “lectivas” trabajando en la tesis, sabiendo que Juan estaba perfectamente cuidado y feliz.

A todos los becarios, que día a día convivimos y “sobrevivimos” a nuestras tesis convirtiéndonos en una gran familia. Vosotros habéis compartido mi día a día muy de cerca, mi boda, mis embarazos, mis hijos, todo ello regado con la salsa de la tesis, que sólo los que la estamos haciendo llegamos a comprender. Gracias a Katha, por su empatía, porque me ha enseñado a que se puede vivir sin carne – se la echa de menos pero se puede ;) – y por sus cursos intensivos de genética. A Quique, porque la sala de becarios no es lo mismo sin él, porque me ha hecho reír todos los días por enseñarme a los wombats… A Natalia, por las eternas charlas "de todos los colores" como dice ella, hablar en el despacho era como escribir un diario =); a Gregor, por cuidarnos a todos, por estar allí cada mañana cuando entraba – imposible ganarle - y por "endulzarnos" la tesis =). A Isa, porque hemos ido en paralelo, siempre de la mano por este camino, me encanta que ahora también terminemos juntas! A Rose, porque es un ejemplo de valentía y de esfuerzo, por su filosofía de vida despreocupada pero a pesar de todo responsable ¡Gracias! A Javi por todos los desayunos (cuando yo aún desayunaba) ¡cuántas historias! A Marina, por su ejemplo, ¡vaya trabajadora! y su dulzura, siempre con una sonrisa. A Paloma, pepito grillo, por ser mi memoria externa, a Ruth, por ser mi compi de al lado, siempre de apoyo, a Andrés, Dani, Nerea, Jesús, Mario, Lauras…en fin, podría hacer otra tesis entera hablando de los despachos!

A mis amigas, con las que ya he compartido más de media vida, Bea, por estar conmigo en todos los momentos importantes ¡hasta ayudó a mis hijos a venir al mundo ¡cómo crecemos!, a Lara, por ser un ejemplo de valentía, por ser mi "medicina" ;), por su visión sencilla de la vida, deberíamos aprender de ti, a Mer, un ejemplo de capacidad de trabajo y de tesón, ¡vas a conseguir todo lo que te propongas!, a María, por vivir a mi lado y hacer la vida más fácil, por ofrecer su ayuda antes de que ni sepa que la necesito, a Je, porque me encanta compartir tiempo con ella, por preguntarme a menudo ¿te he dicho alguna vez..? ;), a Jose y a Di, por las risas, las conversaciones, los viajes, por vuestro tiempo en el Valle del Nansa, realmente se disfruta más la vida a vuestro lado, a Esterilla, por su dulzura y su firmeza, por salvarnos la vida tantas veces, porque en el fondo la gusta el queso y lo sabemos ;) y porque algún día próximo la subiremos al tejado…A Sandra, Chu, Ele…A todas, muchísimas gracias por vuestro cariño con mis peques, por ser una vía de escape en los momentos malos y unas personas geniales con las que compartir los buenos! Espero que la vida nos mantenga juntas mucho más tiempo.

A toda mi nueva familia, Inés y Matías, por ser ejemplo de que siempre se puede seguir adelante, a Almu y Dani porque transmiten su felicidad allá donde van, a César y Paloma por su esfuerzo siempre para tenernos a todos felices. A mis tres preciosos sobris, Lucas Lluvia y Emma, por ser geniales y tan buenos primos para Juan y Mar.

Quiero dar las gracias a mi abuela, por todo su interés en mi tesis. Poca gente habrá seguido tan de cerca mi trabajo. Porque tiene a Dios aburrido de tanto rezar para que todo me vaya bien, porque es la persona con más entusiasmo que conozco y porque ojalá consiga ser como ella. A mi madre, por parecerse a su madre, que de tal palo tal astilla ;) por tener ocho manos todas a mi disposición, por estar dispuesta a cualquier cosa en cualquier momento y todo con una sonrisa, porque es una persona increíble…A mi padre, por enseñarme a valorar mi trabajo, por su gran papel como abuelo, por ser un ejemplo de que querer es poder y lograr poco a poco metas complicadas. A mi hermano, por su alegría, su sensibilidad, su optimismo y su valentía, por ofrecernos su casa y por venir a vernos, ojalá algún día vivas más cerca. A mis niños, Juan y Mar, porque han sido mi inspiración, y la fuerza que me ha impulsado a seguir adelante. Porque arreglan los días grises con una sonrisa, un beso e incluso una travesura.

Y cómo no, a mi compañero inseparable en el camino, por su admiración por mi trabajo y su increíble confianza en mí. Por todas sus preguntas, que me han hecho buscar respuestas. Por vivir a mi lado. Por quererme “infinito elevado a infinito” y porque una vida se hace corta a tu lado…

Index

Figure & Table Index

Figures

Tables

Figures

Tables

Figures

Tables

Figures

Tables

Figures

Tables

Figures in Supporting Information Chapter 1

Tables in Supporting information Chapter 1

Figures in Supporting Information Chapter 2

Tables in Supporting information Chapter 2

Figures in Supporting Information Chapter 3

Figures in Supporting Information Chapter 4

Tables in Supporting Information Chapter 4

Abstract

Climate change is threatening forest ecosystems, driving the need to develop tools to assess species' vulnerability, in order to support conservation management decisions. Maintaining adaptive capacity should be a key target for conservation managers given its importance for increasing species' ability to cope with climate change.

Species distribution models (SDMs) provide a useful tool for conservation by assessing the relationship between species' distribution and environmental variables, and because they can be projected in any geographical space or time. Still, there are some issues that limit the applicability of SDMs in conservation. First of all, until now, SDMs rarely incorporate genetic information, an essential component of adaptive capacity and biodiversity. This trend even worsens in the case of forests trees. Thus, SDMs consider that all populations within a species would respond equally to climatic changes, although the contrary pattern has been reported. Second, there are few examples in literature in which SDMs consider biotic interactions when evaluating vulnerability to climate change. This issue is of large importance as changes in climatic conditions are likely to derive in alterations in species' interactions. Third, most of SDM-based approaches fail in dealing with future uncertainties derived from the wide range of future climate models and scenarios currently available. The main objectives of the present thesis are to overcome these limitations by integrating genetic information, biotic exposure and future climate uncertainties in SDMs, in order to optimize their application for forest management and conservation.

As a first step, we tested the effect of incorporating molecular information in SDMs and we analyzed the relationship among genetic, environmental and geographic distances. We used as a model species maritime pine (*Pinus pinaster* Ait.), for which we defined eight genetically based clades. We compared two SDMs incorporating or not genetic information, and we calculated genetic (Nei's distances), environmental (Euclidean distances in an environmental space) and geographic (least-cost path distances) distances. We found that SDMs improved their predictions when incorporating molecular information and that environment had played a role in shaping the pine's lower order phylogenetic structure.

Second, we evaluated the vulnerability of two iconic Mediterranean species, maritime pine and Aleppo pine (*Pinus halepensis* Mill.), considering their adaptive capacity and their exposure to climate change. We fit SDMs incorporating genetic information, projected them to 42 different future climate predictions corresponding to 2050 in order to consider future climate uncertainties and assessed exposure depending on the concordance among the predictions. Exposure evaluation was then employed to provide specific conservation guidelines for each genetically-defined clade.

Third, we incorporated biotic exposure into vulnerability assessment. We employed as model species *Pinus pinaster* and its potentially important disease, pitch canker, caused by the fungus *Fusarium circinatum* Nirenberg & O'Donnell. We fit SDMs for both, the pine and the disease, we projected them to 18 future climate predictions and evaluated the exposure of the pine to abiotic (climatic) and biotic (pitch canker) exposure. This approach highlighted threatened populations while distinguishing the source of the threat.

Finally, we employed SDMs to provide a standardized methodology to assess marginality considering processes affecting gene flow. Thus, we accounted separately for geographic and environmental factors, to consider isolation by distance and isolation by environment processes respectively. To assess geographic marginality, we used four different indices assessing the position of populations within their core patches and their position in respect to the entire distribution of the species. To assess environmental marginality, we employed the probability of environmental suitability obtained by fitting SDMs. We illustrated our approach with seven European conifers (four of them Mediterranean), namely, *Pinus halepensis* Mill., *Pinus pinaster* Ait., *Pinus nigra* Arnold, *Pinus pinea* L., *Abies alba* Mill., *Pinus sylvestris* L., and *Picea abies* L., all them genetically characterized, in order to perform a second stage analysis, in which we will assess the consequences of marginality on genetic patterns.

Along this thesis, we successfully overcome major limitations of SDMs for their use in conservation management, providing approaches of large applicability in forestry.

Resumen

La gran amenaza que supone el cambio climático para los sistemas forestales hace necesario desarrollar herramientas que evalúen la vulnerabilidad de las especies que los componen, de cara a guiar la gestión para la conservación. Mantener la capacidad adaptativa de las especies debe ser un objetivo principal para los gestores, dada su importancia para lidiar con las alteraciones climáticas.

En el ámbito de la conservación, los modelos de distribución de especies (MDE) son una herramienta útil ya que relacionan la distribución de las especies con variables ambientales pudiendo proyectarse en cualquier marco temporal o geográfico. Sin embargo, hay determinados factores que limitan la aplicabilidad de estos modelos. En primer lugar, hasta ahora, los MDE rara vez incorporan información genética, especialmente en el caso de árboles forestales, considerando, a pesar de las evidencias de lo contrario, que todas las poblaciones de una misma especie reaccionan de igual manera ante los cambios climáticos. En segundo lugar, son escasos los ejemplos en los que los MDE consideran factores bióticos de cara a evaluar la vulnerabilidad de las especies al cambio climático, lo cual es importante ya que los cambios en las condiciones climáticas pueden alterar las interacciones entre especies. Por último, la mayoría de estas aproximaciones no incorporan en sus predicciones la incertidumbre derivada de la gran variedad de modelos y escenarios de clima futuro disponibles. Esta tesis, por tanto, pretende integrar estos elementos (información genética, exposición biótica e incertidumbre climática futura) a lo largo de sus distintos capítulos persiguiendo optimizar la aplicación de los MDE para la gestión forestal y para su conservación.

Como primer paso, testamos el efecto de incorporar información molecular en los MDE y analizamos la relación entre la distancia genética, ambiental y geográfica. Para este objetivo, empleamos como especie modelo el pino marítimo (*Pinus pinaster* Ait.) compuesto por ocho grupos infra-específicos definidos a partir de criterios genéticos. En este trabajo, comparamos dos modelos de nicho (incorporando y no incorporando información genética) y calculamos distancias genéticas (distancia de Nei), ambientales (distancia euclídea en un espacio ambiental) y geográficas (distancias de coste). Como resultado, encontramos una mejora significativa en las predicciones de los modelos que incorporaban información genética así como un papel destacado del ambiente en definir la estructura filogeográfica del pino.

En segundo lugar, evaluamos la vulnerabilidad al cambio climático de dos emblemáticas coníferas Mediterráneas, el pino marítimo y el pino carrasco *(Pinus halepensis* Mill.) teniendo en cuenta su capacidad adaptativa, su exposición al cambio climático y la incertidumbre derivada de las proyecciones climáticas. Para ello, calibramos MDE incorporando información genética, los proyectamos a 42 predicciones climáticas diferentes correspondientes a 2050, de cara a incorporar la incertidumbre derivada de los escenarios climáticos futuros y evaluamos la exposición de las especies dependiendo de la concordancia entre las 42 proyecciones. Posteriormente, utilizamos esta evaluación para plantear directrices de conservación específicas (*in situ*, *ex situ* o *in situ* con monitorización) para cada grupo infra-específico de ambas especies.

En tercer lugar, incorporamos en la evaluación de vulnerabilidad al cambio climático la exposición a factores bióticos. Para ello, usamos como especie modelo *Pinus pinaster* y una de sus potenciales enfermedades, el chancro resinoso, causado por el hongo *Fusarium circinatum* Nirenberg & O'Donnell. Este estudio se desarrolló a nivel nacional (España) de cara a obtener resultados con alta calidad y resolución a partir de los datos disponibles para este territorio. Calibramos MDE para el pino y la enfermedad, proyectamos a 18 escenarios futuros representativos de 2050 y evaluamos la exposición del pino a factores abióticos (climáticos) y bióticos permitiendo detectar poblaciones amenazadas a la vez distinguiendo el factor principal de riesgo.

Finalmente, empleamos los MDEs para definir la marginalidad de una manera estandarizada y considerando los principales procesos que afectan al flujo genético. Para ello, se analizaron de manera independiente factores geográficos y ambientales, de cara a considerar los procesos de aislamiento por distancia y aislamiento por ambiente. Para evaluar la marginalidad geográfica, se definieron cuatro índices caracterizando la posición de las poblaciones dentro de la mancha de distribución en la que encuentran así como su posición con respecto al resto de la distribución de la especie. Para evaluar la marginalidad ambiental, se empleó la probabilidad de adecuación ambiental estimada a través de MDEs. Como especies modelos para el desarrollo de este trabajo, se emplearon siete coníferas Europeas (cuatro de ellas Mediterráneas): *Pinus halepensis* Mill., *Pinus pinaster* Ait., *Pinus nigra* Arnold, *Pinus pinea* L., *Pinus sylvestris* L., *Picea abies* L. y *Abies alba* Mill., todas ellas caracterizadas genéticamente de cara a continuar el estudio con una segunda fase en la que evaluaremos las consecuencias de la marginalidad en los patrones de diversidad genética.

A lo largo de este trabajo, se solucionan con éxito las principales limitaciones de los MDE para su uso en la gestión para la conservación, permitiendo desarrollar aproximaciones de gran aplicación para el mundo forestal.

Chapter 0

General Introduction and Methodological Overview

General Introduction

Forests ecosystems, climate change and conservation

The importance of forest ecosystems is worldwide recognized as they occupy one third of the terrestrial surface, they sustain a myriad of species and represent a livelihood for over 1,600 million of human beings (FAO, 2006, 2010; see Fig. 1 for designated functions of forests). Forests play an essential role as biodiversity hotspots (Myers *et al.*, 2000) given that they smooth regional climate characteristics (*e.g.* López-Carrasco *et al.*, 2015), providing humidity, shadow, refuge and food for a large list of plant, animal and fungi species. As such, 12% of the world's forest cover has been primarily designated for the conservation of biodiversity (FAO, 2010). Forests are also essential in the economy of many regions (see FAO, 2006; Hanewinkel *et al.*, 2012 for global and European data respectively) as a lot of highly demanded products like wood, resin, food and medicine are directly derived from them. In addition, forests ecosystems play an irreplaceable role in protecting water, avoiding soil erosion and maintaining traditional landscapes as well as providing recreational areas. Finally, forest trees are the "green lungs" of the planet as, by means of the photosynthesis, they constitute an oxygen source, which converts them in an essential component for life on Earth. During the photosynthesis, they also become a carbon sink and thus they are a key element for the mitigation of anthropogenic climate change (Canadell & Raupach, 2008).

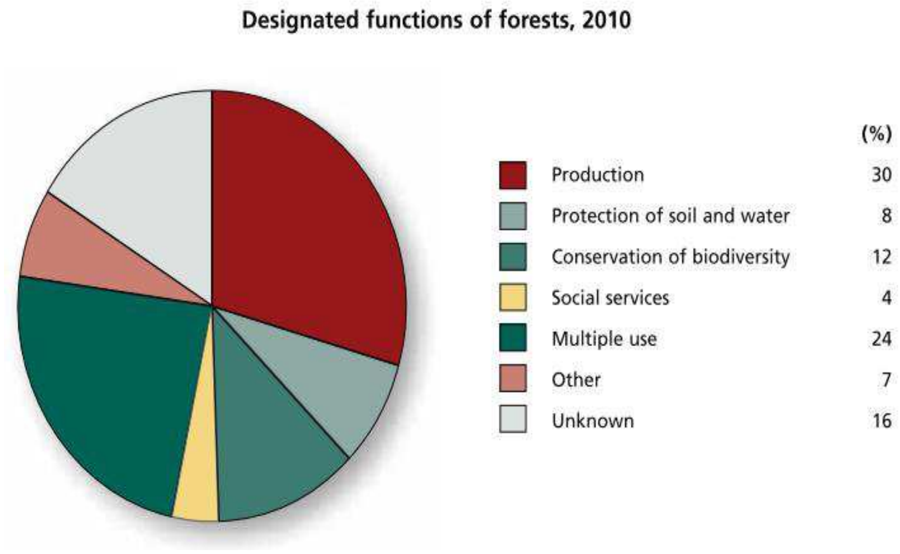


Figure 1: Designations of global forest cover. Figure from FAO, (2010)

Despite their importance, we have lost approximately 1.3 % of the total forest area during the last decade, and although deforestation rates are decreasing, they are still high (data for the period 2000-2010; FAO, 2014). Nevertheless, fortunately, in some regions, such as Europe, we find an inverse trend with an increasing forest cover (Rigo *et al.*, 2016a). In Europe, 33 % of the total land area (215 million ha) are covered by forests from which more than half are coniferous, the rest being broadleaved and mixed (Rigo *et al.*, 2016a; see Fig. 2). Among these, Mediterranean forests located in the Mediterranean Basin, stand out due to their considerably high plant diversity as a result of a noteworthy variety of habitats - *e.g.* 290 wooden species *vs* only 135 for non-Mediterranean Europe -, and of the many historical and paleo-geographic episodes in the area, especially during the last glaciation period. Mediterranean forests are dominated by evergreen species – although deciduous species are also represented - and in particular Mediterranean conifers are characterized by higher within-species diversity than other conifers (Fady, 2005). Accordingly, the Mediterranean Basin, which shelters the vast majority of Mediterranean forests in the world, has been identified as a biodiversity hotspot (Myers *et al.*, 2000).

Anthropogenic climate change, majorly characterized by global warming (IPCC, 2013), is becoming a major threat for natural forests (Thomas *et al.*, 2004; see Allen *et al.*, 2010 for an example) and biodiversity. In the face of climate change, species can migrate, adapt, or become extinct (Aitken *et al.*, 2008) and, in such context, forest ecosystems are especially vulnerable, due to their sessile nature that constrains migration and to their long life-span which does not allow for rapid adaptation to environmental changes (Lindner *et al.*, 2010). In the leading edge of the distribution, migration constitutes the most important process, as trees become main sources of propagules for new available habitats. In contrast, in the trailing edge, adaptive responses of trees are particularly important (Kramer *et al.*, 2010), as it is where species truly face the need to persist in current sites while the environmental conditions are changing (Thuiller *et al.*, 2008). The extent to which populations will adapt, depends on genetic diversity, phenotypic variation (*i.e.* the ability of an individual to change its phenotype responding to environment), strength of selection, fecundity, interspecific competition and biotic interactions (Aitken *et al.*, 2008). Although phenotypic plasticity plays a major role for survival in the short term, evolutionary adaptation becomes crucial in long periods (Thuiller *et al.*, 2008).

Mediterranean regions are particularly vulnerable to climate change (IPCC, 2007; Lindner *et al.*, 2010), due to their position at the rear edge of the distribution of species (Hampe & Petit, 2005), and to the predicted increased frequency of extreme events such as droughts and fires (IPCC, 2007). This threat is particularly relevant not only because of their ecological importance, but also because these forests play an essential role for the society (Gauquelin *et al.*, 2016) – as such, the Mediterranean Basin is considered as an important priority for conservation. Nevertheless, despite their threatened situation, Mediterranean forests remain underrepresented in the current European conservation network (Lefèvre *et al.*, 2013) and in currently available conservation literature (*e.g.* Schueler *et al.*, 2014; but see Noce *et al.*, 2016).

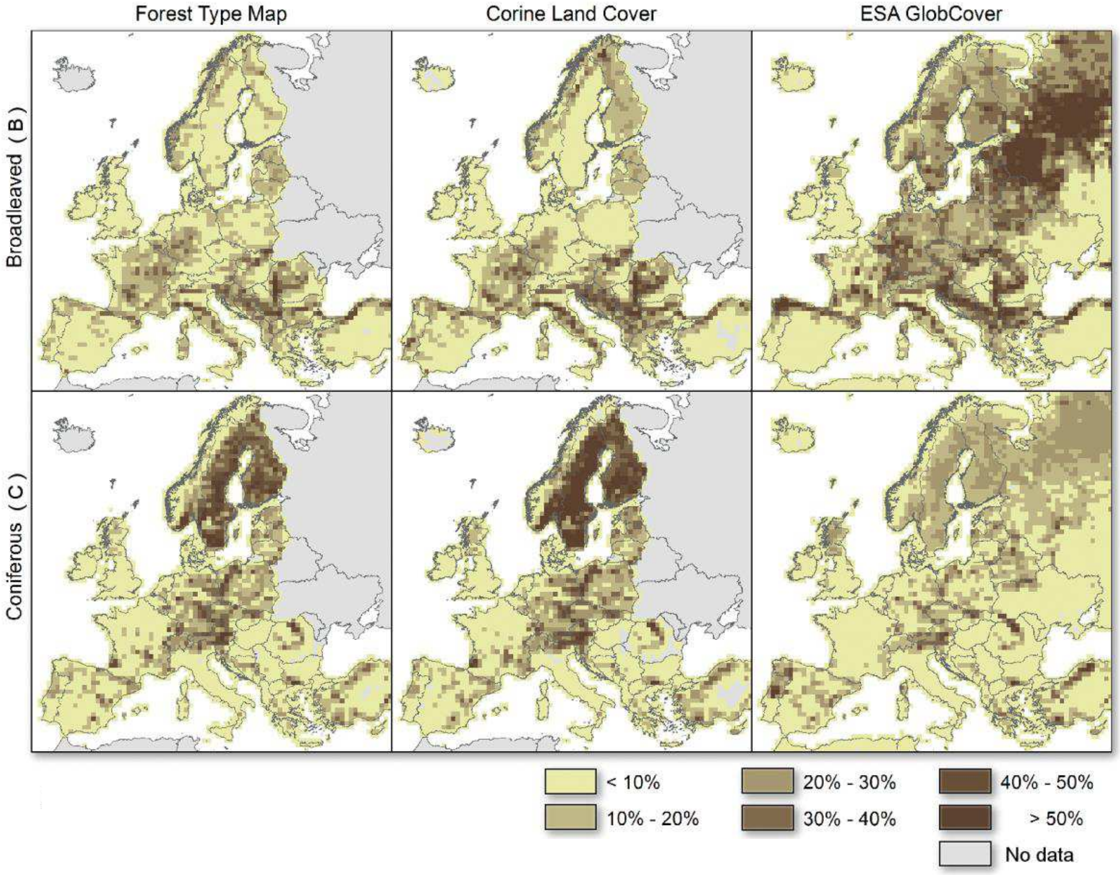


Figure 2: Broadleaved and coniferous forest density, computed with a spatial grid of 50 km^2, for the datasets used: Pan European Forest Type Map 2006 (FTM), CORINE Land Cover map 2006 (CLC) and ESA GlobCover 2009 (EGC). Figure from Rigo *et al.*, (2016b).

Conservation of biodiversity at broad scales is challenging and requires international collaboration to standardize concepts and procedures. Initially, the United Nations Environment Programme (UNEP) gathered experts on biological diversity in 1988, resulting in the development of the convention on biological diversity text (CBD, 1992). This document highlights that conserving biodiversity requires maintaining diversity within species, between species, and between ecosystems. Thus, the CBD extended the goal of conservation from preserving species and their habitats to maintaining their capacity to evolve and adapt to new environmental conditions. In fact, the CBD explicitly highlights the importance of maintaining infra-specific differentiation, and particularly genetic variation as the basis of species divergence when aiming to conserve biodiversity.

Infra-specific differences within forest populations appear due to different processes. Plants rely on pollen and seeds to disperse, but their dispersal abilities are often limited. Thus, when a new factor, such as an environmental or topographic change, appears it may lead to population fragmentation, and consequently to the interruption of gene flow. Within this context, populations evolve independently through neutral and/or adaptive genetic processes resulting in different genetic lineages or clades, an effect that is increased by genetic drift in small populations. If this process continues through time it can ultimately lead to speciation (see Fig. 3).

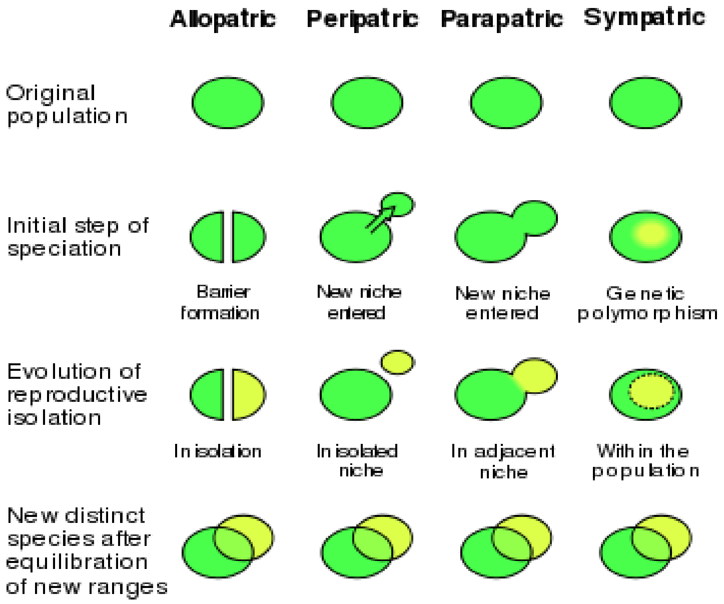


Figure 3 : Population differentiation processes that may ultimately lead to speciation. Figure from Dr. Dana Krempels (included in Lecture notes for evolution and biodiversity -BIL 160 Section HJ at University of Miami).

These genetically differentiated clades, which may show (or not) morphological differences, are likely to diverge in their evolutionary potential and adaptive capacity in which genetic diversity plays a major role (Nicotra *et al.*, 2015). Higher genetic variation implies higher evolutionary potential (Sgró *et al.*, 2011) as selection acts on it, promoting best adapted genotypes and eradicating deleterious ones, ultimately leading to local adaptation. Thus, maintaining or increasing genetic diversity is a major challenge for scientists and managers in the current climatic change context leading to the development of conservation genetics.

The term conservation genetics, was coined more than thirty years ago (Soulé & Wilcox, 1980), and represents the branch of science aiming to maintain genetic diversity through the characterization of population genetic variation, population viability, and future evolution of the species (Frankham, 2010). Evaluating these components enables the delineation of conservation units (CUs), to base conservation plans upon. CUs, all together, should ideally represent the evolutionary potential and heritage of the target species (Moritz, 1994), which depends on the species' genetic and phenotypic variation.

The definition of CUs should rely on empirical information integrating both neutral and adaptive patterns (Fraser & Bernatchez, 2001; Rodríguez-Quilón *et al.*, 2016). Still, the arduousness and economic costs associated to obtain the required field sampled information, has commonly led to the use of alternative strategies. Particularly, in the case of forest trees, eco-geographical zonation has been used as an indicator of homogeneous adaptive genetic zones (*e.g.* Hamann *et al.*, 2004; Schueler *et al.*, 2013) rather than using real genetic measurements (Ledig, 1986), as environment has been proved to be related to genetic variation (*e.g.* Jaramillo-Correa *et al.*, 2015). In fact, in EUFORGEN, the most important international effort for conserving forest tree genetic resources in Europe, the selection of target focal populations (so called dynamic conservation units – DCUs) is based on the stratification of the European environment developed by Metzger *et al.*, (2005). However, but for some exceptions (*e.g.* Vandergast *et al.*, 2008; Zonneveld *et al.*, 2012; Schueler *et al.*, 2013), genetic variation, is still largely neglected in conservation policies (Laikre *et al.*, 2010) although the already visible effects of the warming climate, are leading to an increasing development of conservation plans (*e.g.* Kremen *et al.*, 2008; Gummer *et al.*, 2010) and seed transfer guidelines (e.g Wang *et al.*, 2006; Bower & Aitken, 2008).

While adaptive capacity is essential to select target populations that altogether capture the evolutionary potential of species, assessing the magnitude of the impacts that climate change may have on species *i.e.* their exposure to climate change *sensu* Dawson *et al.*, (2011), is of paramount importance to define conservation strategies and the urgency of the actuations. Once conservation targets have been defined, genetic resources can be maintained with two different strategies: *ex situ* and *in situ* (Ledig, 1986). While the former contemplates a plethora of actions, from the storage of tissues, pollen or life material in gene banks to the plantation of specific conservation stands, *in situ* conservation techniques pursue the preservation of genetic resources in their original ecosystem. Nowadays, *in situ* conservation is the preferred approach as it involves several advantages over *ex situ* conservation (Rotach, 2005): i) it allows for dynamic conservation *i.e.* it preserves natural selection processes, thus maintaining the evolutionary potential of populations ii) it conserves the function of the ecosystem as a whole —not only the target species, but also its associated living organisms— and iii) it is easier, safer and cheaper than *ex situ* conservation. Still, to assess whether *in situ* conservation is possible requires a solid assessment of exposure. However, when *ex situ* conservation is the only option, both, translocation or assisted migration (Leech *et al.*, 2011) and conservation in germplasm banks are possible, the former requiring an assessment of possible locations where it is likely to be successful.

Assessing exposure to climate change requires not only considering abiotic factors, such as climate, but also biotic ones, such as pests, as climate change impacts both abiotic and biotic factors. The predicted impacts of climate change on the formers, such as global warming, or increased intensity and duration of droughts, have been described in detail (*e.g.* IPCC, 2013), and their effects on forests ecosystems have been commonly studied (*e.g.* Coops & Waring, 2011; Schueler *et al.*, 2014). However, the impact of climate change on biotic factors is commonly neglected when assessing exposure to climate change (but see Heikkinen *et al.*, 2007; Preston *et al.*, 2008), although biotic interactions also play an important role in determining the distribution of species (Moore *et al.*, 2007), particularly in a climatic (Araújo & Luoto, 2007; Sturrock *et al.*, 2011) and global (Pautasso *et al.*, 2010) change scenario, as climate change may not only affect species performances, but also species interactions (Tylianakis *et al.*, 2008). For instance, new habitats may become available for predators or pests, or trees may be more prone to infections due to the effects of abiotic factors. Indeed, forest diseases caused by exotic pathogen invasions have been increasing in number and severity (Santini *et al.*, 2013). In addition, studies addressing climate change exposure do not account with future climate uncertainties derived from the wide range of future climatic predictions available, none better than the other (but see Fordham *et al.*, 2011). Thus, future climate uncertainty becomes a major problem when assessing exposure.

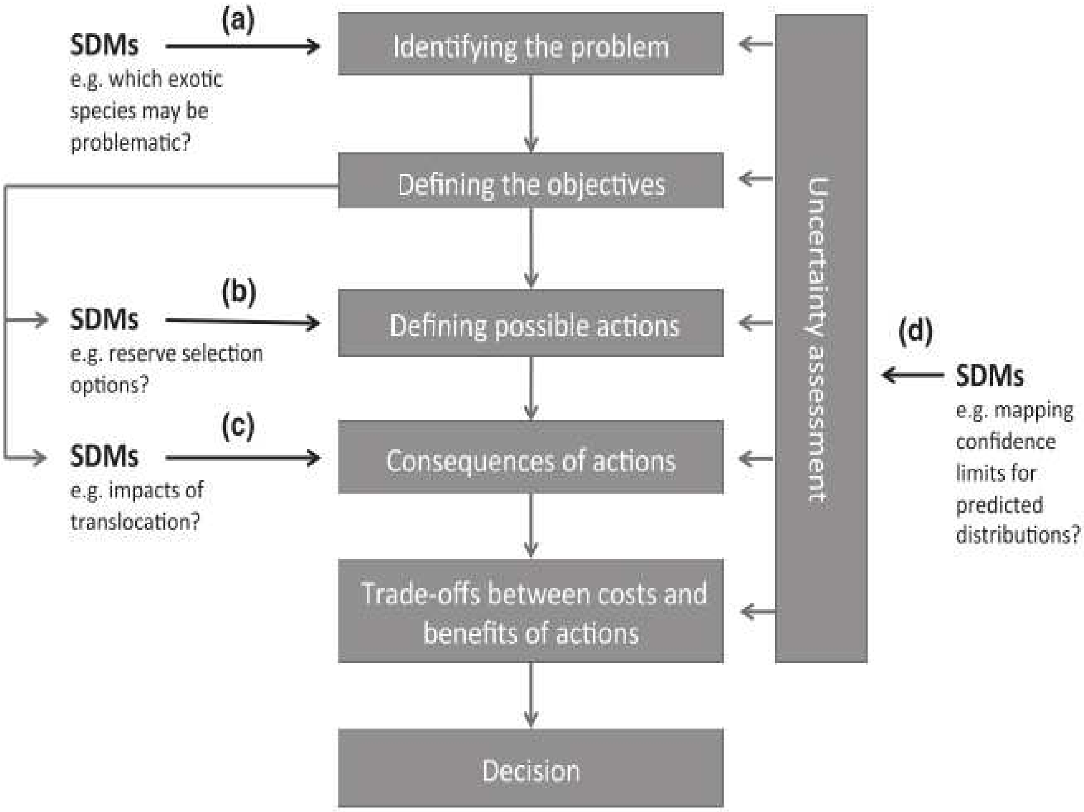


Figure 4: A structured decision-making process with indication of potential entry points for the use of SDMs for conservation. Figure from Guisan *et al.*, (2013)

Throughout this section, we have exposed the new challenges that conservation management is facing. To successfully resolve them, it is necessary to use novel methods combining different branches of science, principally genetics and landscape ecology. Species distribution models (SDMs - see Guisan & Zimmermann, 2000 for details), provide a common framework to develop these kind of approaches, and constitute powerful tools to guide conservation management (see Fig. 4; Guisan *et* *al.*, 2013). These models utilize associations between environmental variables and known species' occurrence records to identify environmental conditions within which populations can be maintained (Pearson, 2010), and their underlying concepts can be easily transferred to the infra-specific level.

Their applications for conservation management are multiple. For instance, they can facilitate our understanding of past processes shaping current genetic structure, which is ultimately interesting to predict future trends, or they make possible an evaluation of exposure to both abiotic and biotic factors. In addition, they enable to apply the concepts and information obtained from genetic/genomics science in a spatially explicit framework, contributing to connect scientists and managers worlds (Shafer *et al.*, 2015). However, despite the high potential SDMs may have in conservational issues, a big effort needs to be undertaken to enhance communication and strengthen collaboration among scientists and managers, as there are few examples in literature in which SDMs support solutions for on-ground conservation problems (Guisan *et al.*, 2013). Because of the high relevance in conservation of SDMs, we will dedicate the next section to introduce their main characteristics.

Species Distribution Models

The origin

Species distribution models (SDMs) evaluate species' habitat suitability based on known occurrence records and on environmental variables. They rely on the well-known concept that ecological factors determine the distribution of species (Grinnell, 1904), and their origin during XXth century, aimed to disentangle how species responded to ecological changes and which were the most important factors driving their distribution patterns (see Grinnell, 1916; Elton, 1927 for examples). As reviewed in Guisan & Thuiller (2005), the earliest examples of SDMs were developped by Johnstons (1924) who predicted the invasive spread of a cactus species in Australia, and Hittinka (1963) who assessed the climatic determinants of the distribution of several European species.

Since the seventies computer-based predictions became an essential tool for the evolution of SDMs, a good example being the work developped by Nix *et al.*, (1977). From then on, the appearance of SDMs in literature increased rapidly (see Ferrier, 1984; Verner *et al.*, 1986; Busby, 1991 for examples; Guisan & Zimmermann, 2000 for a review) including all kind of taxa, from mammals to invertebrates and plants.

During their "growing-up" process, SDMs evolved from a non-spatial, statistical quantification of species-environment relationship based on empirical data or expert knowledge, to an spatially explicit statistical and empirical modeling of species distribution. Objectives also evolved from exclusively explanatory aims (*e.g.* Parra-Olea *et al.*, 2005; Sánchez de Dios *et al.*, 2009), to more prognostic or practical approaches. Nowadays, they are multifunctional tools that are commonly utilized in a plethora of applications, such as evaluating potential invasive species (Broennimann *et al.*, 2007), predicting future distributions in a changing climate framework (Skov & Svenning, 2004), estimating species richness (Steinmann *et al.*, 2009), locating new sample areas for rare species (Guisan *et al.*, 2006), guiding conservation management issues (Marini *et al.*, 2009), simulating palaeo-distributions (Lorenzen *et al.*, 2011) or testing alternative hypotheses of relevant aspects of species history combining SDMs with coalescent models (Richards *et al.*, 2007) or with demographic models (Knowles & Alvarado-Serrano, 2010; Brown & Knowles, 2012).

The modern concept of niche modeling became possible when two parallel streams of research activity converged. On the one hand, field-based ecological studies, enriched with new statistical methods such as Generalized Linear Models (GLM - Nelder & Wedderburn, 1972), Generalized Additive Models (GAM - Hastie & Tibshirani, 1990) or Machine Learning methods (Dietterich, 1996), replaced the previous and more simple regression-based techniques. On the other hand, the huge increase of information availability concerning geography, environmental issues and the development of Geographic Information System as a tool to handle it (Elith & Leathwick, 2009), enabled to characterize the territory.

Concepts and assumptions

First insights on SDMs require a brief summary of the concept of niche, commonly used as a synonymous of habitat or environment. The misunderstanding of these concepts has led to significant confusion about what is actually being modeled and about how the models should be applied (Kearney, 2006). Following Kearney's (2006) work, we will briefly highlight the differences among them.

Habitat refers to the physical and biotical features that are thought to be important for the distribution of an organism, although it can be defined without a reference to the organism *e.g.* savannah or grassland. Moreover, an environment is composed by the biotic and abiotic factors surrounding an organism and potentially interacting with it.

Organisms can modify and even create their own environment (*e.g.* the roots of a plant change soil properties) and thus, although not all components of the organisms' environment will necessarily influence its capacity to grow, survive and reproduce, a reference to the particular organism must be done. To illustrate this concept, in the same habitat two different organisms can experience different environments (*e.g.* diurnal and nocturnal species). Finally, a niche is defined by all the environmental conditions that affect a particular organism, limiting its survival and reproduction. The niche is defined by the organism since its specific properties determine which environmental conditions are relevant.

Following Hutchinson (1957), the niche is conceptualized as a an environmental n-dimensional space whose axes comprise the conditions and resources that limit an organism's survival and reproduction. The idea is similar to the geographical space (two or three dimensional depending on whether we include elevation or not), but in a more complex multi-dimensional space. Hutchinson also established a distinction between fundamental niche (when the effects of biotic interactions are excluded and only abiotic factors are considered) and realized niche (when biotic interactions, *e.g.* competition and predation are included).

The concept of realized niche was broadened later to include all the geographical and historical constraints that influence the distribution of species given their limited ability to reach or re-occupy all suitable areas (see Fig. 5).

Accordingly, significant factors relevant in species distribution patterns are divided into three groups (Guisan & Thuiller, 2005; Soberón & Peterson, 2005): (i) Spatial distribution of environmental conditions favorable to the establishment, survivorship and reproduction of species, (ii) Biotic environment constituted by the species' competitors, predators and pathogens together with the availability and dynamics of resources and, (iii) Species' dispersal capacities. Based on the group of factors selected to define a niche, we can distinguish between two different niche classes: Grinellian and Eltonian (Soberón, 2007). Grinellian niche class is defined fundamentally with non-interactive variables -without the need of resorting to models of exploitation or density competitors- while Eltonian niche class focuses on biotic interactions and resource-consuming dynamics.

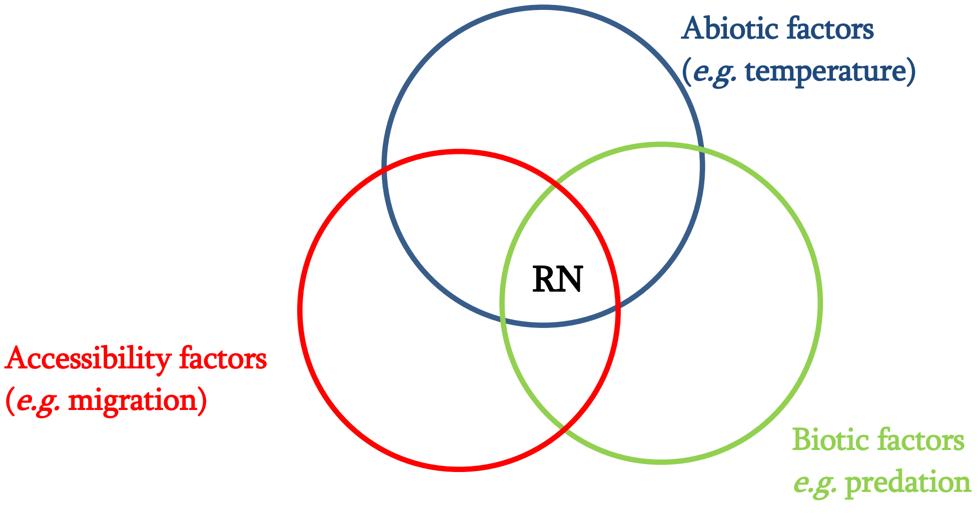


Figure 5 : Realized niche (RN) according to Hutchinson, (1957)

Therefore, Eltonian niche concept coincides with Hutchinson's realized niche idea explained above. Grinellian and Eltonian niches seem to apply to differential spatial scales (global and local scale respectively). Still, the Grinellian niche is the most commonly modeled in SDMs (but see Araújo & Luoto, 2007; Heikkinen *et al.*, 2007). However, because of the observed distribution being already constrained by biotic interactions and limiting resources, SDMs *de facto* quantify Hutchinson's realized niche of species (Guisan & Thuiller, 2005).

Furthermore, when using SDMs, two assumptions are implicitly accepted. First, we assume that the modeled species is in equilibrium with current environmental conditions, *i.e.* it occurs in all suitable areas while being absent from all unsuitable areas. This is unlikely to be true as equilibrium strongly depends on the dispersal ability of the species as well as on biotic interactions, and as sampling is performed during a limited period of time and thus it only reflects a snapshot view (Pearson, 2010). The second premise is that observed occurrence records provide a sample of the entire environmental space occupied by the species, which is not guaranteed by an intense sampling on the geographical space (Pearson, 2010). Although the non-compliance of these two premises leads to restrictions in the use of SDM, there is still very interesting information that can be extracted from them as long as we consider it during the interpretation of the results.

How do SDMs work?

The relationship between environmental variables and species' occurrence records can be addressed through correlative or mechanistic approaches (see Dormann *et al.*, 2012).

While correlative models aim to estimate the environmental conditions that are suitable for a species by associating its known occurrence records with relevant environmental variables that are expected to affect its distribution, mechanistic models pursue the integration of the physiological limiting mechanisms that constrain species' distribution (Pearson, 2010) and their outputs generally consist on fitness maps. Particularly in the case of plant distribution modeling, the bulk of recent publications dealing with SDMs has commonly ignored physiological factors due to the difficulty associated to scaling up from physiological attributes to ecosystem level processes (Higgins *et al.*, 2012) - although some examples can be found in Kearney & Porter (2004, 2009), Crozier & Dwyer (2006) and Morin *et al.* (2008).

In addition, although selecting between a correlative or mechanistic approach depends on the objective, on the spatial scale and commonly on the available information, comparisons between both have reported better predictions from correlative methods as these aim to obtain better performance, while mechanistic approaches pursue understanding the consequences of physiological processes (Kramer *et al.*, 2010). Finally, correlative approaches require more easily obtainable data sets, especially when working at broad scales.

From a conceptual perspective, correlative approaches link the presence/absence of a species (dependent variable) with environmental factors (independent variables) in the environmental space by means of an algorithm and then, it is possible to project the results in the geographic space (see Fig. 6).

There are a wide range of available methods such as regression based - among which GLM and GAM are the most commonly utilized (see Fleishman *et al.*, 2003; Elith & Leathwick, 2007 for examples), machine learning methods – such as Random Forest (RF - Breiman, 2001; see Lorena *et al.*, 2011 for a review and comparison of machine learning classifiers), classification methods like classification tree analysis (CTA; Breiman *et al.*, 1984), or maximum entropy models (*e.g.* MaxEnt; Phillips et al., 2006) among others. While the number of techniques is large and increasing steadily, the discrepancies between them can be very large making it difficult to select the most appropriate methodology (Thuiller et al., 2009).

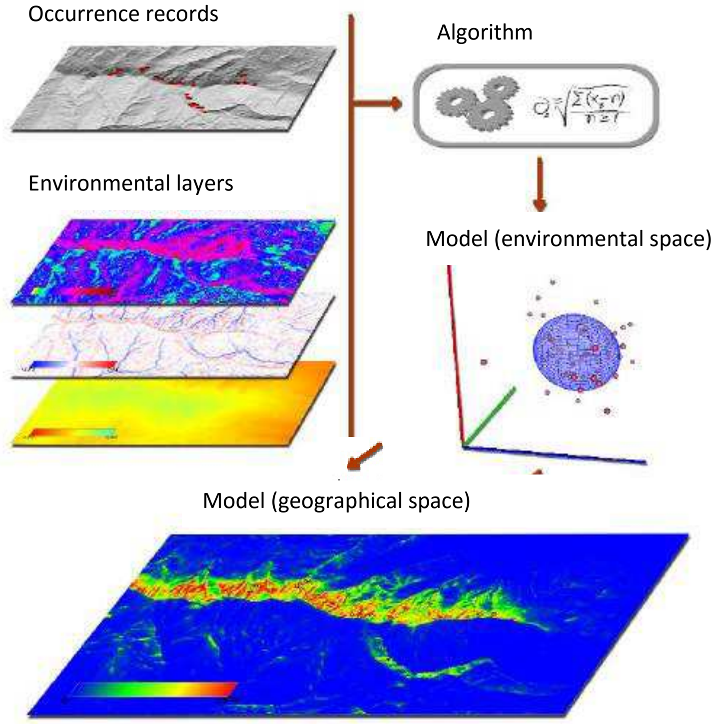


Figure 6 : Conceptual representation of a correlative species distribution model

Several studies have compared the accuracy of these methods (see Olden & Jackson, 2002; Segurado & Araújo, 2004; Elith *et al.*, 2006; Lorena *et al.*, 2011 for examples). However, they have not reached a *consensus* on the superiority of any of them. Currently, there is an increasing use of ensembles of forecasts combining different algorithms (Araújo & New, 2007; Grenouillet *et al.*, 2011) as their use is facilitated by available platforms such as BIOMOD (Thuiller *et al.*, 2009), which provide a common framework to test all the different techniques. Yet, the importance of individual predictions should not be neglected as ensembles remain dependent on them (Araújo & New, 2007).

Data availability is an important issue when selecting the modeling algorithm, which can be categorized in two groups: methods that only require presence data and methods that require presence and absence data (Brotons et al., 2004). Contrary to popular belief, there are very few presence-only SDM algorithms, the most common being rectilinear envelopes (*e.g.* BIOCLIM; Busby, 1991) and distance based envelopes (*e.g.* Mahalanobis distance Farber & Kadmon, 2003). In addition, presence/absence methods have been reported to perform better than presence-only models (Elith *et al.,* 2006) and thus they are the preferred option. However, absence records are scarce and problematic since they might have been recorded when the species was not detected although the environmental variables were suitable for it (false absence). Thus it is very common to use artificially created absence datasets called pseudo-absences or background data. Although these two concepts are very similar, they are not exactly the same. Pseudo-absences records do not include occurrence localities as their aim is to assess differences between occurrence localities and localities chosen to be used in place of real-absence data. Contrarily, background records include occurrence localities as they focus on how the environment where the species is known to occur relates to the environment across the rest of the area (Pearson, 2010). A guidance for the different choices to select pseudo-absences or background records - how, where and how many – is provided in Barbet-Massin *et al.*, (2012).

Whether the environmental variables are categorical or continuous is also a factor to take into account when selecting a modeling algorithm, as not all of them permit the use of the former (Pearson, 2010). Nevertheless, this is not a major task given that a categorical variable can normally be converted into a continuous variable. Finally, whether it is important or not to assess the relative influence of the different environmental factors on the model's fit or predictive capacity is another consideration when selecting the model algorithm (Pearson, 2010). Some models have good predictive power although they do not enable us to understand the relative contributions of different variables.

Other relevant issues when working with SDMs

When working with SDMs, the selection of an appropiate algorithm is relevant as it is the core of the distribution model. Still, it is just one part of the broader modeling process in which a lot more of important decisions must be performed such as the environmental factors to be considered, the spatial scale or the method to evaluate or validate the model.

- Environmental factors

  Determining the factors underlying species' distribution patterns has been a major task in species distribution modeling. However, although there is a wide range of studies concerning the optimization of different statistical techniques and comparing their quality (see previous section), there are comparably few studies that attempt to optimize the selection of model predictors (but see Mac Nally, 2000; Austin *et al.*, 2006). The number of factors to be included in a SDM is still a matter of discussion between scientists, as complex models, which include numerous factors, are more realistic than simpler ones but not necessarily more accurate (Hijmans & Graham, 2006). Finally, it is generally accepted that complex models are likely to be more accurate at finer spatial resolutions or in cases where the species has still not reached equilibrium, whereas simpler models are likely to offer useful and parsimonious solutions at broader scales (Thuiller *et al.*, 2008). Only those factors with direct physiological impact on the species must be selected, as those with indirect effects can cause erroneous predictions when models are used in new regions or under alternative climate scenarios (Guisan & Thuiller, 2005). Currently, remote sensing has provided a new source of information with new environmental datasets. In addition, the use of information extracted from global climate models can provide predictions about the past or the future climate providing new scenarios to project the model.

- Spatial Scale

  The spatial scale, defined by the grain (resolution) and the extent of the study area (Wiens *et al.*, 2002), is a key factor in SDMs as the sample structure and the final results obtained are dependent on it (Dungan *et al.*, 2002; Graham & Hijmans, 2006). It is essential to understand the theory and processes driving the observed distribution patterns to avoid mismatches between the scale used for modeling and the one at which key processes occur (Guisan & Thuiller, 2005). Although conceptually there is not a single natural scale at which ecological patterns should be studied (Levin, 1992), some evidence (see Mackey & Lindenmayer, 2001; Holmes *et al.*, 2005 for examples) suggests that many non-interactive variables (variables considered in Grinellian niche, see Concepts and assumptions section) have broad spatial structures, whereas biotic interactions and resource-consumer dynamics (corresponding to Eltonian niche class) tend to have more fined grained structures (Soberón, 2007). The gradual distribution observed over a large extent and at coarse resolution is likely to be controlled by climate regulators, whereas the patchy distribution observed over a smaller area and at fine resolution is more likely to result from a patchy distribution of resources driven by a micro-topographic variation of habitat fragmentation (Guisan & Thuiller, 2005).

Available information plays a major role when deciding on the spatial scale, as the spatial characteristics of the data, for both environmental (*e.g.* grid cell size) and occurrence records (*e.g.* sample size or geographical accuracy), constrain the decision. The accuracy of the data employed is of paramount importance to obtain reliable results (*e.g.* Bedia *et al.*, 2013). A list of useful information sources to be applied in SDMs of can be found in Carstens & Richards, (2007); Pearson, (2010) and Thomassen *et al.*, (2010).

Concerning climatic information, there are high resolution data sets available that cover large extensions such as WORLDCLIM (Hijmans *et al.*, 2005) for the entire globe or E.OBS (Moreno & Hasenauer, 2016) for Europe. These are of large use and constitute an impressive effort as they compile data sets from different continents and countries. However, they must be utilized with caution, particularly in those areas that lack enough observational points and that are therefore majorly based on interpolations. National meteorological networks provide with more locally accurate datasets such as the the Spanish Meteorological Agency (AEMET) with over 5000 meteorological stations available.

On the other hand, concerning species' distribution data, there are global occurrences data sets, such as GBIF (www.gbif.es), in which researchers worldwide include their observational points of species from all different kingdoms. In particular, for the case of forest trees, and at the European level, there are two main sources of information: (i) the Joint Research Center (JRC - ftp://mars.jrc.ec.europa.eu/Afoludata/Public/DS66/), which encompasses national forestry inventories from different countries and combined them with other sources of information such as satellite imaginary and European ICP plots, to provide high resolution (1 Km) distribution maps (see Köble & Seufert, 2001) and, (ii) EUFORGEN (http://www.euforgen.org/distribution_maps.html), whose distribution maps (polygons shapefiles) are based on expert knowledge and strictly delimitate the native distribution of their target species. As in the case of climatic data, there are local sources of information, such as National Forest Inventories. In Spain, the third Spanish Forest Inventory has already been published and the work for the forth is ongoing.

Independently of the scale selected, SDMs are typically grid-based assuming that environmental variables are homogeneous within each cell. As a result, although broad-scale spatial variability is well represented, small-scale heterogeneity, which can be crucially important for predicting species responses to climate change, is generally not considered (Thuiller *et al.*, 2008).

- Model evaluation

Assessing the robustness and performance of SDMs is challenging. Models with low goodness of fit could successfully explain the relationship between the included variables and the distribution, although there may be other non-considered factors relevant for the distribution of the species. On the opposite, models that perfectly fit the distribution of the species may be over fitted.

There are several statistical techniques to evaluate models performance by comparing predictions with observations (see Fig. 7). These can be based on totally independent evaluation data set, or on resampled observations within the training set, as in the case of cross-validation or bootstrapping. In particular, the area under the ROC curve (AUC - Fielding & Bell, 1997) distinguishes as the most commonly used. Still, its reported problematic issues (see Lobo *et al.*, 2008) have led to the utilization of other alternative metrics such as the true skill statistic (TSS - Allouche *et al.*, 2006) or the H-Measure (Hand, 2010, 2012). In fact, multiple assessments, based on several measures, should be preferred over reporting of a single measure (Fielding, 2002).

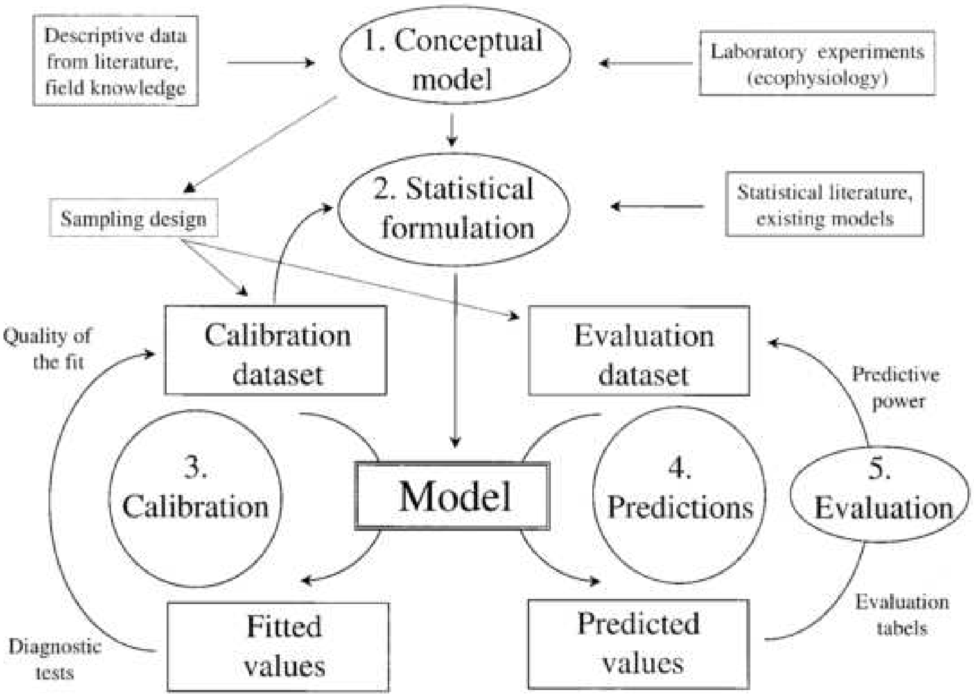


Figure 7: General structure of a correlative species distribution model including evaluation procedure. Figure from Guisan & Zimmermann, (2000)

When dealing with presence-absence models, two types of errors are possible. Commission errors arise from predicting a species where it does not occur, while omission errors stem from failing to predict a species where it does occur.

Molecular information and SDMs

The incorporation of genetic variation in conservation plans has been extensively endorsed (see "Forests ecosystems, climate change and conservation section), as well as the relevance of SDMs in supporting conservation management decisions. Still, very few SDMs integrate infra-specific information, and even less molecular infra-specific information. This fact is surprising given the many authors claiming the importance of combining genetic and environmental variables in a spatially explicit framework (Thuiller *et al.*, 2008; Matyas *et al.*, 2009; Thomassen *et al.*, 2010; Schoville *et al.*, 2012), and given the different responses shown by infra-species groups in their response to climatic changes (Beierkuhnlein *et al.*, 2011).

Although there are some models that incorporate infra-specific information, they are generally based on sub-species (*e.g.* Pearman *et al.*, 2010; Oney *et al.*, 2013), and those based on molecular markers are, up to now, exclusively for animals (Pease *et al.*, 2009; D'Amen *et al.*, 2013). In the case of plants, O'Neill *et al.* (2008) used an universal transfer function based on different species provenances tested in multi-site trials to model the distribution and productivity of lodgepole pine. In addition, Benito-Garzón *et al.*, (2011) integrated infra-specific information in SDMs to calibrate models of tree survival for maritime and scots pines using adaptive traits quantified from provenance trials. Both studies highlighted the importance of considering infra-specific information when modelling species' responses to climate change, as future predicted distributions varied largely depending on the populations modelled. Nevertheless, to our knowledge, integrating genetic information directly derived from molecular markers (for instance by delineating genetically defined clades) has still not been addressed in the literature. Consequently, here, we integrated already available infra-specific information, directly derived from molecular markers (majorly from Single Nucleotide Polymorphisms – SNPs) into SDMs aiming to provide support for conservation management.

Methodological overview

Structure of the thesis

This thesis aims to enhance the applicability of SDMs in conservation management by integrating molecular information and biotic interactions while incorporating future climate uncertainties.

We focused on Mediterranean conifers and developed multi-specific approaches including taxa with different evolutionary histories to ensure that our approach is valid independently of the target.

We used correlative SDMs (see "How do SDMs work" for justification) to model the distribution of our target species majorly based on exclusively climate-related variables *i.e.* abiotic factors. Still, we implicitly modelled the realized niche of species as our datasets, which were based on real occurrences, reflected all the factors (abiotic, biotic and dispersal related) that constrained the distribution of the target species. In this thesis, we explicitly incorporated biotic interactions, in order to provide a framework to consider these in SDMs which is particularly important when assessing climate change impacts. Finally, when projecting the obtained SDMs into future climate scenarios, we used a wide range of future climate predictions to deal with future climate uncertainties.

We organized the work in subsequent chapters as follows (see Table 1):

Chapter 1 addresses two objectives: (i) to test whether integrating molecular information into SDMs improved their predictions and, (ii) to disentangle the role of environment and geography on delineating the infra-specific structure of *Pinus pinaster* Ait (maritime pine). We found that molecular information improved SDMs' predictions and that environment played a role in defining maritime pine's lower order phylogenetic structure.

Chapter 2 assesses vulnerability by means of two of its major components, *i.e.* adaptive capacity and exposure, using SDMs incorporating molecular information. We illustrated our approach with two iconic Mediterranean conifers – *P. pinaster* and *Pinus halepensis* Mill. (Aleppo pine). Based on the exposure component of vulnerability, we recommended optimal conservation strategies distinguishing among *in situ*, *in situ* with monitoring and *ex situ* conservation strategies and we evaluated current European efforts for the conservation of forest genetic resources.

In Chapter 3, we incorporated biotic interactions into SDMs. We analyzed the exposure to climate change of maritime pine considering both, abiotic (climate) and biotic (pitch canker disease caused by the fungus *Fusarium circinatum* Nirenberg & O'Donnell) factors. We aimed to provide an accurate and high resolution tool for the management of the species and thus, we performed this study at a national scale (Spain) for which we had information fulfilling these requirements. We recommended setting up breeding programmes in highly exposed and productive populations, while silvicultural methods and monitoring should be applied in those less productive but still exposed populations.

Finally, in Chapter 4, we provided a standardized methodology to delineate marginal populations considering two relevant processes for gene flow, isolation by distance and isolation by environment. Thus, we defined geographic and environmental indices (the former based on geometry and the latter on SDMs predictions) to assess both aspects of marginality. This work set the basis for a second-stage analysis including molecular information, in which we will assess the effects of marginality on genetic patterns. We illustrated our approach with four Mediterranean conifers, *i.e. Pinus halepensis* Mill., *Pinus pinaster* Ait., *Pinus nigra* Arnold, *Pinus pinea* L., and three other European conifers *Pinus sylvestris* L., *Picea abies* L. and *Abies alba* Mill., all them genetically characterized. This chapter describes on going work, and thereby, we provide preliminary results and discussion.

In this work we have used WORLDCLIM, EUFORGEN and JRC data for those studies at the European level (Chapter 1, Chapter 2, Chapter 4), and the Spanish national forestry inventory and Spanish Meteorological information (AEMET) when performing the analysis at a national (Spain) level (Chapter 3).

Concerning genetic information, we compiled data from all available studies covering the (almost) entire distribution of the target species to delineate genetically defined clades. For *Pinus pinaster*, genetically defined clades were based on maternally inherited markers (mitochondrial DNA sequences - Burban & Petit, 2003), paternally inherited markers (chloroplast Simple Sequence Repeats - SSRs; Vendramin *et al.*, 1998; Bucci *et al.*, 2007) and bi-parental markers namely 12 nuclear Simple Sequence Repeats (nSSRs) and 266 Single Nucleotide Polymorphisms

(SNPs), the latter being a subset of the 384 SNPs from 772 individuals/36 populations analysed in Jaramillo-Correa *et al.*, (2015). In the case of *Pinus halepensis*, we delineated genetic clades based on a SNP dataset (1325 individuals from 49 populations) published in Serra-Varela *et al.* (in revision).

Table 1: Overview of the structure of this thesis, including objectives, spatial extent, target species, algorithms and results in forms of manuscripts.

	OBJECTIVE	SPATIAL EXTENT	MODEL SPECIES	ALGORITHMS	RESULTS
Chapter 1	(i) Influence of molecular information in Species Distribution Models (ii) Role of environment and geography on infra-specific structure	Western Mediterranean Basin and European Atlantic coast.	*Pinus pinaster* Ait.	Generalized Linear Models (GLM) Generalized Additive Models (GAM) Random Forest (RF) Classification tree Analysis (CTA) MaxEnt	Serra-Varela *et al.*, (2015); Global Ecology and Biogeography
Chapter 2	(i) Vulnerability assessment (ii) Selection of adequate conservation strategies based on exposure	Europe	*Pinus pinaster* Ait. *Pinus halepensis* Mill.	Generalized Linear Models (GLM) Generalized Additive Models (GAM) Random Forest (RF) Classification tree Analysis (CTA) MaxEnt	Serra-Varela *et al.*, in revision in Diversity and Distributions
Chapter 3	(i) Incorporate biotic exposure in SDMs (ii) Provide a useful tool for forest management	Spanish Iberian Peninsula	*Pinus pinaster* Ait.	Generalized Additive Models (GAM)	Serra-Varela *et al.* in revision in PLOSONE
Chapter 4	(i) Delineation of marginal populations from a geographical and environmental perspective (ii) Set the basis to analyze marginal populations from a genetic perspective.	Europe	*Pinus pinaster* Ait. *Pinus halepensis* Mill. *Pinus nigra* Arnold. *Pinus pinea* L. *Pinus sylvestris* L. *Abies alba* Mill. *Picea abies* L.	Generalized Additive Models (GAM)	Serra-Varela *et al.*, in prep

References

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.

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

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.

Araújo, M.B. & Luoto, M. (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, **16**, 743–753.

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, **22**, 42–47.

Austin, M.P., Belbin, L., Meyers, J.A., Doherty, M.D. & Luoto, M. (2006) Evaluation of statistical models used for predicting plant species distributions: Role of artificial data and theory. *Ecological Modelling*, **199**, 197–216.

Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **3**, 327–338.

Bedia, J., Herrera, S., Gutiérrez, J.M. & Manuel, J. (2013) Dangers of using global bioclimatic datasets for ecological niche modeling. Limitations for future climate projections. *Global and Planetary Change*, 1–12.

Beierkuhnlein, C., Thiel, D., Jentsch, A., Willner, E. & Kreyling, J. (2011) Ecotypes of European grass species respond differently to warming and extreme drought. *Journal of Ecology*, **99**, 703–713.

Benito-Garzón, M., Alía, R., Robson, T.M. & Zavala, M.A. (2011) Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography*, **20**, 766–778.

Bower, A.D. & Aitken, S.N. (2008) Ecological genetics and seed transfer guidelines for *Pinus albicaulis* (*Pinaceae*). *American journal of botany*, **95**, 66–76.

Breiman, L. (2001) Random forests. *Machine Learning*, **45**, 5–32.

Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. (1984) *Classification and Regression Trees*, Wadsworth & Brooks/Cole Advanced Books & Software, Monterey.

Broennimann, O., Treier, U. a, Müller-Schärer, H., Thuiller, W., Peterson, A.T. & Guisan,

A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology letters*, **10**, 701–709.

Brotons, L., Thuiller, W., Araújo, M.B. & Hirzel, A.H. (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, **27**, 437–448.

Brown, J.L. & Knowles, L.L. (2012) Spatially explicit models of dynamic histories: examination of the genetic consequences of Pleistocene glaciation and recent climate change on the American Pika. *Molecular ecology*, **21**, 3757–3775.

Bucci, G., González-Martínez, S.C., Le Provost, G., Plomion, C., Ribeiro, M.M., Sebastiani, F., Alía, R. & Vendramin, G.G. (2007) Range-wide phylogeography and gene zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. *Molecular Ecology*, **16**, 2137–2153.

Burban, C. & Petit, R.J. (2003) Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Molecular Ecology*, **12**, 1487–1495.

Busby, J.R. (1991) *BIOCLIM - a bioclimate analysis and prediction system. Nature Conservation Cost Effective Biological Surveys and Data Analysis* (ed. by C.R. Margules) and M.P. Austin), pp. 64–68. CSIRO.

Canadell, J.G. & Raupach, M.R. (2008) Managing forests for climate change mitigation. *Science*, **320**, 1456.

Carstens, B.C. & Richards, C.L. (2007) Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution; international journal of organic evolution*, **61**, 1439–54.

CBD (1992) *Convention on Biological Diversity Text*, United Nations Environment Programme, Rio de Janeiro.

Coops, N.C. & Waring, R.H. (2011) Estimating the vulnerability of fifteen tree species under changing climate in Northwest North America. *Ecological Modelling*, **222**, 2119–2129.

Crozier, L. & Dwyer, G. (2006) Combining population-dynamic and ecophysiological models to predict climate-induced insect range shifts. *The American naturalist*, **167**, 853–866.

D'Amen, M., Zimmermann, N.E. & Pearman, P.B. (2013) Conservation of phylogeographic lineages under climate change. *Global Ecology and Biogeography*, **22**, 93–104.

Dawson, T.P., Jackson, S.T., House, J.I., Prentice, I.C. & Mace, G.M. (2011) Beyond predictions: biodiversity conservation in a changing climate. *Science*, **332**, 53–58.

Dietterich, T. (1996) Machine learning. *ACM Computing Surveys*, **28**.

Dormann, C.F., Schymanski, S.J., Cabral, J., Chuine, I., Graham, C.H., Hartig, F., Kearney, M.R., Morin, X., Römermann, C., Schröder, B. & Singer, A. (2012) Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, **39**, 2119–2131.

Dungan, J.L., Dale, M.R.T., Legendre, P. & Citron-Pousty, S. (2002) A balanced view of scale in spatial statistical analysis. *Ecography*, **25**, 626–640.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species ' distributions from occurrence data. *Ecography*, **29**, 129–151.

Elith, J. & Leathwick, J. (2007) Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. *Diversity and Distributions*, **13**, 265–275.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 1–19.

Elton, C. (1927) *Animal Ecology*, (ed. by Methuen) Sidgwick and Jackson.

Fady, B. (2005) Is there really more biodiversity in Mediterranean forest ecosystems? *Taxon*, **54**, 905–910.

FAO (2006) *Better forestry, less poverty; A practitioner's guide*, Food and Agriculture Organization of the United Nations, Rome.

FAO (2010) *Global Forest Resources Assessment 2010*, Food and Agriculture Organization of the United Nations, Rome; Italy.

FAO (2014) *State of the World's Forests*, Food and Agriculture Organization of the United Nations, Rome, Italy.

Farber, O. & Kadmon, R. (2003) Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling*, **160**, 115–130.

Ferrier, S. (1984) The Status of the Rufous Scrub-Bird Atrichornis rufescens: habitat geographical variation and abundance.

Fielding, A.H. (2002) *What are the appropriate characteristics of an accuracy measure? Predicting Species Occurrences Issues of Accuracy and Scale* (ed. by J.M. Scott), P.J. Heglund), M.L. Morrison), J.B. Haufler), M.G. Raphael), W.A. Wall), and F.B. Samson), p. 868. Island Press.

Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.

Fleishman, E., Nally, R. Mac & Fay, J.P. (2003) Validation tests of predictive models of butterfly occurrence based on environmental variables. *Conservation Biology*, **17**, 806–817.

Fordham, D.A., Wigley, T.M.L. & Brook, B.W. (2011) Multi-model climate projections for biodiversity risk assessments. *Ecological Applications*, **21**, 3317–3331.

Frankham, R. (2010) Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation*, **143**, 1919–1927.

Fraser, D.. & Bernatchez, L. (2001) Adaptive evolutionary conservation : towards a unified concept for defining coservation units. *Molecular Ecology*, **10**, 2741–2752.

Gauquelin, T., Michon, G., Joffre, R., Duponnois, R., Génin, D., Fady, B., Bou Dagher-Kharrat, M., Derridj, A., Slimani, S., Badri, W., Alifriqui, M., Auclair, L., Simenel, R., Aderghal, M., Baudoin, E., Galiana, A., Prin, Y., Sanguin, H., Fernandez, C. & Baldy, V. (2016) Mediterranean forests, land use and climate change: a social-ecological perspective. *Regional Environmental Change*.

Graham, C.H. & Hijmans, R.J. (2006) A comparison of methods for mapping species ranges and species richness. *Global Ecology and Biogeography*, **15**, 578–587.

Grenouillet, G., Buisson, L., Casajus, N. & Lek, S. (2011) Ensemble modelling of species distribution: The effects of geographical and environmental ranges. *Ecography*, **34**, 9–17.

Grinnell, J. (1916) The niche-relationship of the California Thrasher. *The Auk*, **34**, 427–433.

Grinnell, J. (1904) The origin and distribution of the chestnut-backed chickadee. *The Auk*, **21**, 364–382.

Guisan, A., Broennimann, O., Engler, R., Vust, M., Yoccoz, N.G., Lehmann, A. & Zimmermann, N.E. (2006) Using niche-based models to improve the sampling of rare species. *Conservation Biology*, **20**, 501–511.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S. a, Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology letters*, **16**, 1424–1435.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.

Gummer, D.L., Schumaker, N.H., Bender, D.J. & Heinrichs, J.A. (2010) Assessing critical habitat: Evaluating the relative contribution of habitats to population persistence. *Biological Conservation*, **143**, 2229–2237.

Hamann, A., Aitken, S.N. & Yanchuk, A.D. (2004) Cataloguing *in situ* protection of genetic resources for major commercial forest trees in British Columbia. *Forest Ecology and Management*, **197**, 295–305.

Hampe, A. & Petit, R.J. (2005) Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, **8**, 461–467.

Hand, D.J. (2012) Assessing the performance of classification methods. *International Statistical Review*, **80**, 400–414.

Hand, D.J. (2010) Evaluating diagnostic tests: The area under the ROC curve and the balance of errors. *Statistics in Medicine*, **29**, 1502–1510.

Hanewinkel, M., Cullmann, D.A., Schelhaas, M.J., Nabuurs, G.J. & Zimmermann, N.E. (2012) Climate change may cause severe loss in the economic value of European forest land. *Nature Climate Change*, **3**, 203–207.

Hastie, T.J. & Tibshirani, R. (1990) *Generalized additive models*, Chapman & Hall/CRC, Boca Raton - London - New York - Washington D.C.

Heikkinen, R.K., Luoto, M., Virkkala, R., Pearson, R.G. & Körber, J.H. (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, **16**, 754–763.

Higgins, S.I., O'Hara, R.B., Bykova, O., Cramer, M.D., Chuine, I., Gerstner, E.-M., Hickler, T., Morin, X., Kearney, M.R., Midgley, G.F. & Scheiter, S. (2012) A physiological analogy of the niche for projecting the potential distribution of plants. *Journal of Biogeography*, **39**, 2132–2145.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hijmans, R.J. & Graham, C.H. (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, **12**, 2272–2281.

Holmes, K.W., Kyriakidis, P.C., Chadwick, O. a., Soares, J.V. & Roberts, D. a. (2005) Multi-scale variability in tropical soil nutrients following land-cover change. *Biogeochemistry*, **74**, 173–203.

Hutchinson, G.E. (1957) Concluding remarks- Cold Spring Harbor Symposia on Quantitative Biology. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415–427.

IPCC (2013) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, (ed. by T.F. Stocker, D. Qin, G.K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley) Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,.

IPCC (2007) *Climate Change 2007: impacts, adaptation and vulnerability: contribution of Working Group II to the fourth assessment report of the Intergovernmental Panel*, Cambridge.

Jaramillo-Correa, J.P., Rodríguez-Quilón, I., Grivet, D., Lepoittevin, C., Sebastiani, F., Heuertz, M., Garnier-Géré, P., Alía, R., Plomion, C., Vendramin, G. & González-Martínez, S. (2015) Molecular proxies of climate maladaption in a long-lived tree. *Genetics*, **199**, 1–15.

Kearney, M.R. (2006) Habitat, environment and niche: what are we modelling? *Oikos*, **115**, 186–191.

Kearney, M.R. & Porter, W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology letters*, **12**, 334–50.

Kearney, M.R. & Porter, W.P. (2004) Mapping the fundamental niche: Physiology, climate, and the distribution of a nocturnal lizard. *Ecology*, **85**, 3119–3131.

Knowles, L.L. & Alvarado-Serrano, D.F. (2010) Exploring the population genetic consequences of the colonization process with spatio-temporally explicit models: insights from coupled ecological, demographic and genetic models in montane grasshoppers. *Molecular ecology*, **19**, 3727–3745.

Köble, R. & Seufert, G. (2001) *Novel maps for forest tree species in Europe. Proceedings of the 8th European Symposium on the Physico-Chemical Behaviour of Air Pollutants: "A Changing Atmosphere!,"* pp. 17–20. Data set available online at: http://afoludata.jrc.ec.europa.eu/index.php/dataset/detail/66, Torino.

Kramer, K., Degen, B., Buschbom, J., Hickler, T., Thuiller, W., Sykes, M.T. & de Winter, W. (2010) Modelling exploration of the future of European beech (*Fagus sylvatica* L.) under climate change—Range, abundance, genetic diversity and adaptive response. *Forest Ecology and Management*, **259**, 2213–2222.

Kremen, C., Cameron, A., Moilanen, A., Phillips, S.J., Thomas, C.D., Beentje, H., Dransfield, J., Fisher, B.L., Glaw, F., Good, T.C., Harper, G.J., Hijmans, R.J., Lees, D.C., Louis, E., Nussbaum, R.A., Raxworthy, C.J., Razafimpahanana, A., Schatz, G.E., Vences, M., Vieites, D.R., Wright, P.C. & Zjhra, M.L. (2008) Aligning conservation priorities across taxa in Madagascar with high-resolution planning tools. *Science*, **320**, 222–226.

Laikre, L., Allendorf, F.W., Aroner, L.C., Baker, C.S., Gregovich, D.P., Hansen, M.M., Jackson, J.A., Kendall, K.C., McKelvey, K., Neel, M.C., Olivieri, I., Ryman, N., Schwartz, M.K., Bull, R.S., Stetz, J.B., Tallmon, D.A., Taylor, B.L., Vojta, C.D., Waller, D.M. & Waples, R.S. (2010) Neglect of genetic diversity in implementation of the convention on biological diversity. *Conservation Biology*, **24**, 86–88.

Ledig, E.T. & Ledig, F.T. (1986) Conservation strategies for forest gene resources. *Forest Ecology and Management*, **14**, 77–90.

Leech, S.M., Almuedo, P.L. & Neill, G.O. (2011) Assisted Migration : Adapting forest management to a changing climate. *BC Journal of Ecosystems and Management*, **12**, 18–34.

Lefèvre, F., Koskela, J., Hubert, J., Kraigher, H., Longauer, R., Olrik, D.C., Schueler, S., Bozzano, M., Alizoti, P., Bakys, R., Baldwin, C., Ballian, D., Black-Samuelsson, S., Bednarova, D., Bordács, S., Collin, E., de Cuyper, B., de Vries, S.M.G., Eysteinsson, T., Frýdl, J., Haverkamp, M., Ivankovic, M., Konrad, H., Koziol, C., Maaten, T., Notivol

Paino, E., Oztürk, H., Pandeva, I.D., Parnuta, G., Pilipovič, A., Postolache, D., Ryan, C., Steffenrem, A., Varela, M.C., Vessella, F., Volosyanchuk, R.T., Westergren, M., Wolter, F., Yrjänä, L. & Zariŋa, I. (2013) Dynamic conservation of forest genetic resources in 33 European countries. *Conservation biology: the journal of the Society for Conservation Biology*, **27**, 373–84.

Levin, S.A. (1992) The problem of pattern and scale in ecology. *Ecology*, **73**, 1943–1967.

Lindner, M., Maroschek, M., Netherer, S., Kremer, A., Barbati, A., Garcia-Gonzalo, J., Seidl, R., Delzon, S., Corona, P., Kolstr??m, M., Lexer, M.J. & Marchetti, M. (2010) Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. *Forest Ecology and Management*, **259**, 698–709.

Lobo, J.M., Jiménez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145–151.

López-Carrasco, C., López-Sánchez, A., San Miguel, A. & Roig, S. (2015) The effect of tree cover on the biomass and diversity of the herbaceous layer in a Mediterranean dehesa. *Grass and Forage Science*, **70**, 639–650.

Lorena, A.C., Jacintho, L.F.O., Siqueira, M.F., Giovanni, R. De, Lohmann, L.G., de Carvalho, A.C.P.L.F. & Yamamoto, M. (2011) Comparing machine learning classifiers in potential distribution modelling. *Expert Systems with Applications*, **38**, 5268–5275.

Lorenzen, E.D., Nogués-Bravo, D., Orlando, L., Weinstock, J., Binladen, J., Marske, K.A. & Cooper, A. (2011) Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature*, **479**, 359–364.

Mackey, B.G. & Lindenmayer, D.B. (2001) Towards a hierarchical framework for modelling the spatial distribution of animals. *Journal of Biogeography*, **28**, 1147–1166.

Marini, M.A., Barbet-Massin, M., Esteves Lopes, L. & Jiguet, F. (2009) Major current and future gaps of Brazilian reserves to protect Neotropical savanna birds. *Biological conservation*, **142**, 3039–3050.

Matyas, C., Nagy, L. & Jarmay, E.U. (2009) *Genetic background of response of trees to aridification at the xeric forest limit and consequences for bioclimatic modelling. Bioclimatology and Natural Hazards*, pp. 179–196. Springer, Netherlands.

Metzger, M.J., Bunce, R.G.H., Jongman, R.H.G., Mücher, C. a. & Watkins, J.W. (2005) A climatic stratification of the environment of Europe. *Global Ecology and Biogeography*, **14**, 549–563.

Moore, P., Hawkins, S.J. & Thompson, R.C. (2007) Role of biological habitat amelioration in altering the relative responses of congeneric species to climate change. *Marine Ecology Progress Series*, **334**, 11–19.

Moreno, A. & Hasenauer, H. (2016) Spatial downscaling of European climate data. *International Journal of Climatology*, **36**, 1444–1458.

Morin, X., Viner, D. & Chuine, I. (2008) Tree species range shifts at a continental scale : new predictive insights from a process-based model. *Journal of Ecology*, **96**, 784–794.

Moritz, C. (1994) Defining evolutionarily-significant-units for conservation. *Trends in Ecology & Evolution*, **9**, 373–375.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 72.

Mac Nally, R. (2000) Regression and model-building in conservation biology, biogeography and ecology: The distinction between - and reconciliation of - “predictive” and “explanatory” models. *Biodiversity and Conservation*, **9**, 655–671.

Nelder, J.A. & Wedderburn, R.W.M. (1972) Generalized Linear Models. *Journal of the Royal Statistical Association*, **135**, 370–384.

Nicotra, A.B., Beever, E.A., Robertson, A.L., Hofmann, G.E. & O’Leary, J. (2015) Assessing the components of adaptive capacity to improve conservation and management efforts under global change. *Conservation Biology*, **29**, 1268–1278.

Nix, H.A., McMahon, J. & MacKenzie, M.D. (1977) *Potential areas of production and the future of Pigeon Pea and other grain legumes in Australia. The potential for Pigeon Pea in Australia Proceedings of Pigeon Pea Cajanus cajan L Millsp Field Day* (ed. by E.S. Wallis) and P.C. Whitman), pp. 5/1–5/12. University of Queensland, Australia.

Noce, S., Collalti, A., Valentini, R. & Santini, M. (2016) Hot spot maps of forest presence in the Mediterranean basin. *iForest - Biogeosciences and Forestry*, **809**.

O’Neill, G.A., Hamann, A. & Wang, T. (2008) Accounting for population variation improves estimates of the impact of climate change on species’ growth and distribution. *Journal of Applied Ecology*, **45**, 1040–1049.

Olden, J.D. & Jackson, D.A. (2002) A comparison of statistical approaches for modelling fish species distributions. *Freshwater Biology*, **47**, 1976–1995.

Oney, B., Reineking, B., O’Neill, G.A. & Kreyling, J. (2013) Intraspecific variation buffers projected climate change impacts on *Pinus contorta*. *Ecology and Evolution*, **3**, 437–449.

Parra-Olea, G., Martinez-Meyer, E. & Pérez-Ponce de León, G. (2005) Forecasting climate change effects on salamander distribution in the highlands of central Mexico. *Biotropica*, **37**, 202–208.

Pautasso, M., Dehnen-Schmutz, K., Holdenrieder, O., Pietravalle, S., Salama, N., Jeger, M.J., Lange, E. & Hehl-Lange, S. (2010) Plant health and global change - Some implications for landscape management. *Biological Reviews*, **85**, 729–755.

Pearman, P.B., D’Amen, M., Graham, C.H., Thuiller, W. & Zimmermann, N.E. (2010) Within-taxon niche structure: niche conservatism, divergence and predicted effects of climate change. *Ecography*, **33**, 990–1003.

Pearson, R.G. (2010) Species’ distribution modeling for conservation educators and practitioners. *Lessons in Conservation*, **3**, 54–89.

Pease, K.M., Freedman, A.H., Pollinger, J.P., McCormack, J.E., Buermann, W., Rodzen, J., Banks, J., Meredith, E., Bleich, V.C., Schaefer, R.J., Jones, K. & Wayne, R.K. (2009) Landscape genetics of California mule deer (*Odocoileus hemionus*): the roles of ecological and historical factors in generating differentiation. *Molecular Ecology*, **18**, 1848–1862.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Preston, K.L., Rotenberry, J.T., Redak, R.A. & Allen, M.F. (2008) Habitat shifts of endangered species under altered climate conditions: importance of biotic interactions. *Global change biology.*, **14**, 2501–2515.

Richards, C.L., Carstens, B.C. & Lacey Knowles, L. (2007) Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *Journal of Biogeography*, **34**, 1833–1845.

Rigo, D. De, Bosco, C., Durrant, T.H., San-Miguel-Ayanz, J., Barredo, J.I., Strona, G., Caudullo, G., Leo, M. Di & Boca, R. (2016a) *Forest resources in Europe : an integrated perspective on ecosystem services , disturbances and threats. The European Atlas of Forest Tree Species :*, pp. 8–19.

Rigo, D. De, Caudullo, G., Durrant, T.H. & San-Miguel-Ayanz, J. (2016b) *Modelling , data and information on forest tree species. The European Atlas of Forest Tree Species*, pp. 40–45.

Rodríguez-Quilón, I., Santos-del-Blanco, L., Serra-Varela, M.J., Koskela, J., González-Martínez, S.C. & Alía, R. (2016) Capturing neutral and adaptive genetic diversity for conservation: maritime pine as a case study. *Ecological Applications*, **accepted**.

Rotach, P. (2005) In situ *conservation. Conservation and Management of Forest Genetic Resources in Europe* (ed. by T. Geburek) and J. Turok), pp. 535–565. Arbora Publishers, Zvolen.

Sánchez de Dios, R., Benito-Garzón, M. & Sainz-Ollero, H. (2009) Present and future extension of the Iberian submediterranean territories as determined from the distribution of marcescent oaks. *Plant Ecology*, **204**, 189–205.

Santini, A., Ghelardini, L., De Pace, C., Desprez-Loustau, M.L., Capretti, P., Chandelier, A., Cech, T., Chira, D., Diamandis, S., Gaitniekis, T., Hantula, J., Holdenrieder, O., Jankovsky, L., Jung, T., Jurc, D., Kirisits, T., Kunca, A., Lygis, V., Malecka, M., Marcais, B., Schmitz, S., Schumacher, J., Solheim, H., Solla, A., Szabò, I., Tsopelas, P., Vannini, A., Vettraino, a. M., Webber, J., Woodward, S. & Stenlid, J. (2013) Biogeographical patterns and determinants of invasion by forest pathogens in Europe. *New Phytologist*, **197**, 238–250.

Schoville, S.D., Bonin, A., François, O., Lobreaux, S., Melodelima, C. & Manel, S. (2012) Adaptive Genetic Variation on the Landscape: Methods and Cases. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 23–43.

Schueler, S., Falk, W., Koskela, J., Lefèvre, F., Bozzano, M., Hubert, J., Kraigher, H., Longauer, R. & Olrik, D.C. (2014) Vulnerability of dynamic genetic conservation units of forest trees in Europe to climate change. *Global Change Biology*, **20**, 1498–1511.

Schueler, S., Kapeller, S., Konrad, H., Geburek, T., Mengl, M., Bozzano, M., Koskela, J., Lefèvre, F., Hubert, J., Kraigher, H., Longauer, R. & Olrik, D.C. (2013) Adaptive genetic diversity of trees for forest conservation in a future climate: a case study on Norway spruce in Austria. *Biodiversity and Conservation*, **22**, 1151–1166.

Segurado, P. & Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555–1568.

Serra-Varela, M.J., Alía, R., Pórtoles, J., Gonzalo-Jiménez, J., Soliño, M., Grivet, D. & Raposo, R. (submitted). Incorporating exposure to pitch canker disease to support management decisions of *Pinus pinaster* Ait. in the face of climate change. *Forest Ecology and Management*.

Serra-Varela, M.J., Alía, R., Ruíz-Daniels, R., Zimmermann, N.E., Gonzalo-Jiménez, J. & Grivet, D. (in revision). Assessing vulnerability of two Mediterranean conifers to support genetic conservation management. *Diversity and Distributions*.

Serra-Varela, M.J., Grivet, D., Vincenot, L., Broennimann, O., Gonzalo-Jiménez, J. & Zimmermann, N.E. (2015) Does phylogeographic structure relate to climatic niche divergence? A test using maritime pine (*Pinus pinaster* Ait.). *Global Ecology and Biogeography*, **24**, 1302–1313.

Sgró, C.M., Lowe, A.J. & Hoffmann, A.A. (2011) Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications*, **4**, 326–337.

Shafer, A.B.A., Wolf, J.B.W., Alves, P.C., Bergström, L., Bruford, M.W., Brännström, I., Colling, G., Dalén, L., De Meester, L., Ekblom, R., Fawcett, K.D., Fior, S., Hajibabaei, M., Hill, J.A., Hoezel, A.R., Höglund, J., Jensen, E.L., Krause, J., Kristensen, T.N., Krützen, M., McKay, J.K., Norman, A.J., Ogden, R., Österling, E.M., Ouborg, N.J., Piccolo, J., Popovic, D., Primmer, C.R., Reed, F.A., Roumet, M., Salmona, J., Schenekar, T., Schwartz, M.K., Segelbacher, G., Senn, H., Thaulow, J., Valtonen, M., Veale, A., Vergeer, P., Vijay, N., Vilà, C., Weissensteiner, M., Wennerström, L., Wheat, C.W. & Zielinski, P. (2015) Genomics and the challenging translation into conservation practice. *Trends in Ecology and Evolution*, **30**, 78–87.

Skov, F. & Svenning, J. (2004) Potential impact of climatic change on the distribution of forest herbs in Europe. *Ecography*, **27**, 366–380.

Soberón, J. (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology letters*, **10**, 1115–23.

Soberón, J. & Peterson, A.T. (2005) Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, **2**, 1–10.

Soulé, M.E. & Wilcox, B.M. (1980) *Conservation Biology: An Evolutionary-ecological Perspective.*, Sunderland, USA.

Steinmann, K., Linder, H.P. & Zimmermann, N.E. (2009) Modelling plant species richness using functional groups. *Ecological Modelling*, **220**, 962–967.

Sturrock, R.N., Frankel, S.J., Brown, A. V., Hennon, P.E., Kliejunas, J.T., Lewis, K.J., Worrall, J.J. & Woods, A.J. (2011) Climate change and forest diseases. *Plant Pathology*, **60**, 133–149.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L. J. Collingham, Y.C. & Hughes, L. (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.

Thomassen, H.A., Cheviron, Z.A., Freedman, A.H., Harrigan, R.J., Wayne, R.K. & Smith, T.B. (2010) Spatial modelling and landscape-level approaches for visualizing intra-specific variation. *Molecular Ecology*, **19**, 3532–3548.

Thuiller, W., Albert, C., Araújo, M.B., Berry, P.M., Cabeza, M., Guisan, A., Hickler, T., Midgley, G.F., Paterson, J., Schurr, F.M., Sykes, M.T. & Zimmermann, N.E. (2008) Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 137–152.

Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009) BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369–373.

Tylianakis, J.M., Didham, R.K., Bascompte, J. & Wardle, D.A. (2008) Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, **11**, 1351–1363.

Vandergast, A.G., Bohonak, A.J., Hathaway, S. a., Boys, J. & Fisher, R.N. (2008) Are hotspots of evolutionary potential adequately protected in southern California? *Biological Conservation*, **141**, 1648–1664.

Vendramin, G.G., Anzidei, M., Madaghiele, A. & Bucci, G. (1998) Distribution of genetic diversity in Pinus pinaster Ait. as revealed by chloroplast microsatellites. *Theoretical and Applied Genetics*, **97**, 456–463.

Verner, J., Morrison, M.L. & Ralph, C.J. (1986) *Wildlife 2000: modeling habitat relationships of terrestrial vertebrates*, University of Wisconsin Press, Madison, WI.

Wang, T., Hamann, A., Yanchuk, A.D., O'Neill, G.A. & Aitken, S.N. (2006) Use of response functions in selecting lodgepole pine populations for future climates. *Global Change Biology*, **12**, 2404–2416.

Wiens, J.A., Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A. & Samson, F.B. (2002) *Predicting Species occurences: progress, Problems and prospects. Predicting species occurences issues of accuracy and scale* (ed. by J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Raphael, W.A. Wall and F.B. Samson), pp. 739–749. Island Press.

Zonneveld, M. Van, Scheldeman, X., Escribano, P., Viruel, M. a, Van Damme, P., Garcia, W., Tapia, C., Romero, J., Sigueñas, M. & Hormaza, J.I. (2012) Mapping genetic diversity of cherimoya (Annona cherimola Mill.): application of spatial analysis for conservation and use of plant genetic resources. *PLoS ONE*, **7**, e29845.

Chapter 1

Does phylogeographic structure relate to climatic niche divergence? A test using maritime pine (*Pinus pinaster* Ait.)

Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2015) **24**, 1302–1313

Global Ecology and Biogeography A Journal of Macroecology

RESEARCH PAPERS

Does phylogeographical structure relate to climatic niche divergence? A test using maritime pine (*Pinus pinaster* Ait.)

M. J. Serra-Varela[1,2,3], D. Grivet[3], L. Vincenot[3,4,5], O. Broennimann[6], J. Gonzalo-Jiménez[1,2]* and N. E. Zimmermann[7]

[1]Department of Plant Production and Forest Resources, University of Valladolid, Avda de Madrid 44, 34004 Palencia, Spain, [2]Sustainable Forest Management Research Institute, INIA, University of Valladolid, 34004 Palencia, Spain, [3]Department of Forest Ecology and Genetics, INIA, Forest Research Centre, 28040 Madrid, Spain, [4]CREAF, Campus Universitat Autonoma de Barcelona, Cerdanyola del Vallès 08193, Spain, [5]Ecodiv – UE1293, Université de Rouen, 76821 Mont-Saint-Aignan, France, [6]Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland, [7]Landscape Dynamics, Swiss Federal Research Institute WSL, CH-8903 Birmensdorf, Switzerland

ABSTRACT

Aim To disentangle the effects of environmental and geographical processes driving phylogenetic distances among clades of maritime pine (*Pinus pinaster*). To assess the implications for conservation management of combining molecular information with species distribution models (SDMs; which predict species distribution based on known occurrence records and on environmental variables).

Location Western Mediterranean Basin and European Atlantic coast.

Methods We undertook two cluster analyses for eight genetically defined pine clades based on climatic niche and genetic similarities. We assessed niche similarity by means of a principal component analysis and Schoener's *D* metric. To calculate genetic similarity, we used the unweighted pair group method with arithmetic mean based on Nei's distance using 266 single nucleotide polymorphisms. We then assessed the contribution of environmental and geographical distances to phylogenetic distance by means of Mantel regression with variance partitioning. Finally, we compared the projection obtained from SDMs fitted from the species level (SDMsp) and composed from the eight clade-level models (SDMcm).

Results Genetically and environmentally defined clusters were identical. Environmental and geographical distances explained 12.6% of the phylogenetic distance variation and, overall, geographical and environmental overlap among clades was low. Large differences were detected between SDMsp and SDMcm (57.75% of disagreement in the areas predicted as suitable).

Main conclusions The genetic structure within the maritime pine subspecies complex is primarily a consequence of its demographic history, as seen by the high proportion of unexplained variation in phylogenetic distances. Nevertheless, our results highlight the contribution of local environmental adaptation in shaping the lower-order, phylogeographical distribution patterns and spatial genetic structure of maritime pine: (1) genetically and environmentally defined clusters are consistent, and (2) environment, rather than geography, explained a higher proportion of variation in phylogenetic distance. SDMs, key tools in conservation management, better characterize the fundamental niche of the species when they include molecular information.

*Correspondence: Julián Gonzalo-Jiménez, Sustainable Forest Management Research Institute, University of Valladolid, Avda de Madrid 44, 34004 Palencia, Spain.
E-mail: jgonzalo@pvs.uva.es

Keywords
Climate change, conservation biology, conservation, genetics, infraspecies, niche conservatism, SDM, species distribution model.

INTRODUCTION

The role of the environment in genetic divergence and whether or not species tend to retain aspects of their fundamental niche more strongly than expected from their phylogenetic relatedness (i.e. niche conservatism versus niche evolution) are still matters of significant debate (Wiens & Graham, 2005). The different criteria used to report whether niches are similar or not have led

DOI: 10.1111/geb.12369

to different conclusions in the literature (Warren *et al.*, 2008). A further question is the degree of overlap between genetic relatedness and environmental niche similarity and, more specifically, how phylogenetic distance among species is correlated to, or explained by, environmental or geographical distances.

These issues relate to major forces of speciation and can be transferred to the infraspecific level, i.e. to the evolutionary significant units (ESUs) that characterize a species (Moritz, 1994). ESUs define sets of populations within a species that are historically distinct and likely to have a distinct evolutionary potential: see Bucci *et al.* (2007), Pearman *et al.* (2010) and D'Amen *et al.* (2013) for examples with maritime pine, herpetofauna and African mammals, respectively. Species evolutionary units play a pivotal role in interpreting a species' history, i.e. colonization, migration or adaptation events (Rodríguez-Sánchez *et al.*, 2010). If different ESUs follow very distinct evolutionary trajectories, speciation can ultimately occur.

Species distribution models (SDMs), when combined with phylogenetic information, allow us to test niche conservatism versus dynamism (Pearman *et al.*, 2008) and to answer the questions we have posed above. While SDMs enable us to characterize environmental niches and to map the areas potentially occupied by a species, phylogenies inform us about evolutionary relatedness among taxa.

SDMs calibrate associations between environmental variables and known species occurrence records to identify the environmental conditions necessary for the viability of a population (Guisan & Thuiller, 2005). This association can be expressed either by correlative or by mechanistic approaches, and models can be based on just abiotic predictors or may also include biotic ones. A common procedure when comparing the environments of two sister species is to first calibrate SDMs for both species separately, and then to evaluate the overlap in their predicted niche distribution (Warren *et al.*, 2008).

Despite the proliferation of SDMs during the last decade, only a limited number of studies have integrated infraspecific information based on morphologically defined subspecies or clades (e.g. Pearman *et al.*, 2010; Oney *et al.*, 2013), and even fewer have done so with information based on genetically defined groups. This is surprising, given that the importance of combining genetic information with environmental variables in a spatially explicit framework has been pointed out by many authors (Thuiller *et al.*, 2008; Thomassen *et al.*, 2010). Integrating molecular information into SDMs seems especially important when projecting these models under future environmental conditions, as genetically defined subspecies or clades have been found to diverge in their niche characteristics when fitted under current climate conditions (Pearman *et al.*, 2010; D'Amen *et al.*, 2013). These findings suggest local adaptation and niche differentiation and, therefore, a different response to climatic changes.

Previous studies investigating ecological differences between genetically based infraspecific clades have focused mainly on animal subspecies (e.g. Pease *et al.*, 2009; D'Amen *et al.*, 2013). In the case of plants, Benito-Garzón *et al.* (2011) integrated infraspecific information to calibrate models of tree survival for maritime and Scots pine using adaptive traits quantified from provenance trials. Similarly, O'Neill *et al.* (2008) modelled the distribution and productivity of lodgepole pine using different species provenances tested in multisite trials. Both studies highlighted the importance of considering the different populations of a species when modelling its responses to climate change, as future predicted distributions varied largely depending on the populations modelled.

None of the previous studies has fully explored the level of environmental space overlap, or whether it is related to the (phylo)genetic similarities among the clades evaluated. Such information is important for determining whether niche adaptation, that can potentially lead to speciation, is a significant force behind the spatial and environmental structure of infraspecific clades. Alternatively, niche conservatism could lead to more conserved ecological niches than would be suggested from genetic or spatial distance between clades. Such information is ultimately relevant for designing conservation plans that integrate locally adapted populations as essential components within the context of climate change.

To sum up, to further understand the phylogenetic, environmental and spatial relationships across the whole range of a species, there is a need for studies integrating infraspecific molecular and ecological information in a phylogeographical framework. Our study, using maritime pine (*Pinus pinaster* Ait.) as a model species, helps to fill this gap, as we calibrate SDMs and combine genetic, environmental and geographical information. This mediterranean pine is ecologically versatile, growing on a variety of substrates (from siliceous to calcareous), across a wide range of elevations (from sea level to 600 m, and up to 2000 m in its southern range in Morocco) and under different climates (semi-arid to humid). Its strong ecological differentiation and patchy distribution has led to the definition of several varieties (Barbéro *et al.*, 1998). In addition, both adaptive phenotypic traits (Alia *et al.*, 1997; Sierra-de-Grado *et al.*, 2008) and molecular resources from all three genomes, namely, mitochondrial, chloroplast and nuclear (Burban & Petit, 2003; Bucci *et al.*, 2007; Santos-del-Blanco *et al.*, 2012; Jaramillo-Correa *et al.*, 2015), analysed across the full distribution range, show spatially structured patterns. In combination, these characteristics suggest that maritime pine populations may be adapted to distinct niches, making this conifer an excellent model for studying local adaptation and differentiation at the infraspecific level.

The present study aims at answering three main questions: (1) does niche overlap in maritime pine correlate with phylogenetic relatedness at the clade level; (2) is the phylogenetic distance among clades better explained by environmental or by geographical distance and (3) do infraspecific SDMs representing distinct evolutionary units provide better fits and spatial projections than species-level SDMs? To answer these questions, we first fitted and validated ensemble SDMs (i.e. sets of several models) for each genetic clade, and for the species fitted as a whole. We also analysed two predicted distributions of the species, one for the species fitted from species-level data and one composed from the clade-level models. Second, we estimated the niche position of each clade along principal components analysis (PCA) axes based on environmental data, and compared genetic, environmental and geographical distances. Finally, we assessed the differences between the species-level and the infraspecific clade-level SDMs regarding predictive performance.

METHODS

Molecular data

The combination of mitochondrial, chloroplast and nuclear (simple sequence repeats, SSRs, and single nucleotide polymorphisms, SNPs) molecular markers from previous studies allowed us to define eight genetic clades for the full distribution range of maritime pine: Atlantic Iberian Peninsula (G1), eastern populations (G2), Atlantic France (G3), Morocco (G4), eastern (G5), central (G6) and southern (G7) Spain, and Tunisia (G8) (Fig. 1) (see Appendix S1 in Supporting Information for a detailed explanation). Three marginal populations (hatched symbols in Fig. 1) were unclassified, as we considered it inappropriate to assign them to any of the eight defined clusters due to insufficient field-sampled molecular information. Another isolated population (Fuencaliente, FCN, Spain), which was genotyped just for SNPs, was excluded from the present study.

We used a priori neutral markers to define the eight clades, i.e. primarily influenced by demographic processes and not by adaptation. While adaptive selection, especially the SNPs, cannot be completely excluded (see Jaramillo-Correa *et al.*, 2015), the fact that each independent set of molecular markers points to the same overall genetic structure strongly supports our selection.

To assess the genetic relatedness among the clusters, Nei's genetic distances (Nei, 1978) were calculated for 266 SNPs (Jaramillo-Correa *et al.*, 2015) and used to build a phylogenetic cladogram with unweighted pair group method with arithmetic mean (UPGMA), with 10,000 bootstrap replications using POPTREE2 (Takezaki *et al.*, 2010).

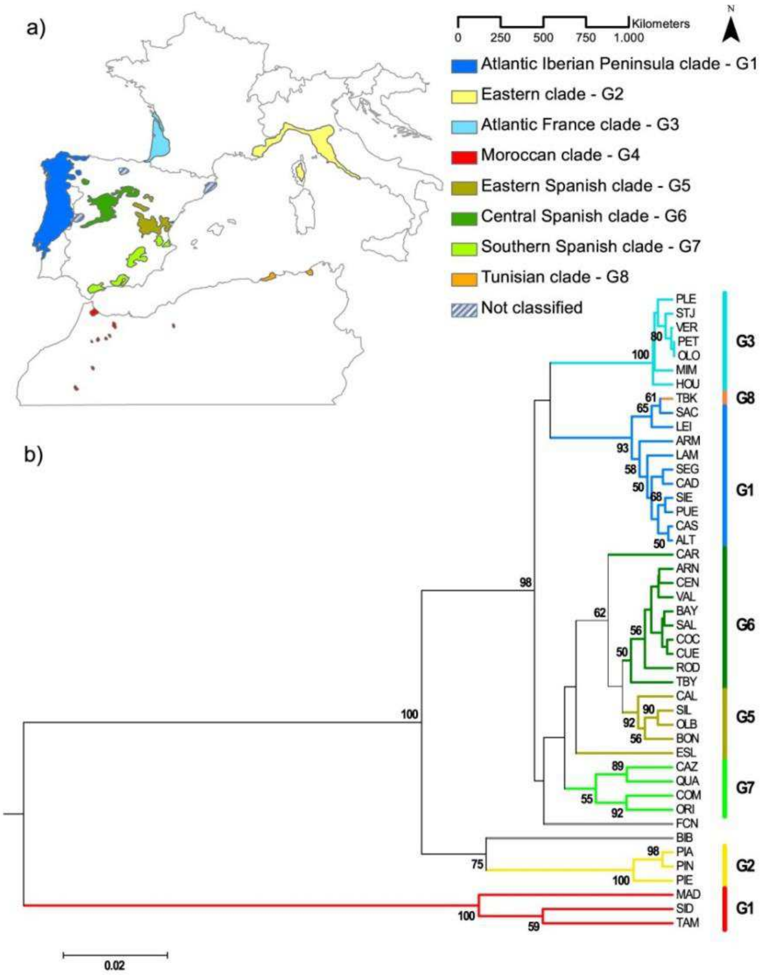


Figure 1 (a) Distribution of genetic clades (G1–G8) along the native distribution range of maritime pine. Hatched areas correspond to non-classified clades. (b) Unweighted pair group method with arithmetic mean (UPGMA) dendrogram based on pairwise population Nei's distances on 266 SNPs in 45 populations, 10,000 bootstrap replications. Only bootstrap support values greater than or equal to 50% are displayed.

Species data

Two different sources of species occurrence were combined to represent the complete native range of maritime pine: (1) the Tree Species Distribution for Europe (TSDE; Köble & Seufert, 2001) from the Joint Research Centre's AFOLU data portal (ftp://mars.jrc.ec.europa.eu/Afoludata/Public/DS66/), and (2) the EUFORGEN database from the European forest genetic resources programme (http://www.euforgen.org/distribution_maps.html). TSDE maps tree species percentage occupancy in Europe at a 1-km grid level, but it does not differentiate between native and planted populations. EUFORGEN delivers a distribution shape that, while spatially less accurate, includes the whole species range within Europe and northern Africa, and specifically excludes planted populations. Therefore, by filtering TSDE occurrences with EUFORGEN, we obtained a good approximation of the native range of maritime pine. Finally, we split the species' distribution range into eight subranges based on our eight, geographically distinct, clades (see Appendix S1 for details on maritime pine clades).

A set of presence and pseudo-absence records was created for the species and for each of the genetically defined infraspecific clades. A total of 128,653 presences were used for the calibration of the model at the species level, and from 59,109 (G1) to 3997 (G8) at the clades level. Pseudo-absences (643,265 records for all the models) were randomly selected within the study area (see Fig. 1b) within those cells where TSDE reported 0% occupancy (see Appendix S2 for more details).

Bioclimatic data

We considered 23 bioclimatic variables representative of the period 1950–2000 for the analysis. Nineteen of them (BIO1–BIO19) were downloaded from the current bioclimatic variables available in WorldClim (Hijmans *et al.*, 2005). Four new variables were created following Zimmermann *et al.* (2007) to better characterize water availability: summer and spring potential evapotranspiration ($ETPT_{summer}$, $ETPT_{spring}$) and summer and spring moisture index ($MIND_{summer}$, $MIND_{spring}$) (see Appendix S3 for a detailed explanation of bioclimatic data).

After screening for correlation and variance inflation, we selected four climate variables as predictors for our analyses: BIO3 (isothermality), BIO11 (mean temperature of the coldest quarter), BIO13 (precipitation of the wettest month) and $MIND_{summer}$ (see Appendix S3 for a detailed explanation of variable selection).

Species distribution models

We used five different statistical methods to calibrate the models individually, namely general linear models (GLMs; McCullagh & Nelder, 1989), generalized additive models (GAMs; Hastie & Tibshirani, 1990), random forest (RF; Breiman, 2001), classification tree analysis (CTA; Breiman *et al.*, 1984) and MaxEnt (Phillips *et al.*, 2006). We built 45 individual algorithm models, five for each of the eight genetic clades, and five for the species level. All models were processed in BIOMOD (Thuiller *et al.*, 2009) using the package 'biomod2' in the R statistical software environment (R Development Core Team, 2013), which was also used for all additional statistical analyses. Model performance was assessed by means of the true skill statistic (TSS), the area under the receiver operating characteristic curve (AUC) and the H-measure (see Appendix S2 for more information on algorithm specifications and model performance assessment).

Several reviews in the past have compared the advantages of using different statistical methods (e.g. Elith *et al.*, 2006; Bedia *et al.*, 2011), but today using just one model or the best among several is considered a less powerful approach than using several models in an ensemble (Araújo & New, 2007; Grenouillet *et al.*, 2011). Thus, we averaged the predictions from the five statistical methods we had used to create nine ensemble models, one for each genetic clade and one for the species level.

The large number of occurrence records available permitted a random division of each dataset (corresponding to the species and to each clade) into two equally sized subsets for training and evaluating. Both subsets maintained the initial proportion between presence and pseudo-absence records. Evaluation was performed on 20% of the set-aside evaluation records (see Appendix S2 for more details).

Probabilistic model outputs based on current climatic conditions were converted to binary presence–absence maps by defining thresholds that optimized TSS values. Thereby, nine different distribution maps were created, eight corresponding to the different genetically defined clade models (G1–G8 projections) and one to the species-level model (referred to as 'SDMsp').

Niche overlap

Niche overlap was determined in geographical and environmental space. In geographical space, projections G1 to G8 were used to create a single map ('overlapping clade projection') highlighting the overlap in bioclimatically suitable areas among the different clades. Values could range from zero (i.e. indicating that bioclimatic conditions are unsuitable for any clade) to eight (i.e. indicating cells that are suitable for all genetic clades). Then, geographical overlap was quantified for every pair of clades as the percentage suitable for both clades among the total area suitable for any of the two clades.

To measure niche overlap in environmental space we applied the methodology proposed by Broennimann *et al.* (2012). This recent technique corrects for spatial autocorrelation in the environment, thus reducing the possibility of confounding effects between spatial patterns and ecological processes (Warren *et al.*, 2014). The method uses Schoener's D metric (Schoener, 1970) as a measure of environmental niche overlap, and includes a statistical framework to test for niche similarity and equivalency, as proposed by Warren *et al.* (2008). The idea is to create a multivariate environmental grid, based on the first two axes of a PCA summarizing all the selected environmental variables, in which each cell represents a unique combination of the environmental conditions available in the study area. Then, Gaussian kernel density functions are used to estimate the occupancy of each cell (z_{ij}), and the *D* metric is calculated based on the different z_{ij} values obtained. Finally, equivalency and similarity tests in both directions were performed. To create the PCA axes, we used the same four bioclimatic variables that were used to calibrate the SDMs across all study-area pixels (PCA-env in Broennimann *et al.*, 2012). The median PCA score obtained for the presence records of each genetic clade was used as an indicator of its position along the PCA axes, and the variance among the group as an estimator of niche breadth in each axis.

Constructing SDMs for species and molecular-based clades

In order to create a composite model for the whole species based on the sum of all molecular-based clades, projections G1 to G8 were superimposed. We refer to this result as the 'composite clade model' or SDMcm, as opposed to the 'species-level model' or SDMsp. It differs from the 'overlapping clade model' in that it only separates areas that are climatically suitable and unsuitable for the species. SDMcm and SDMsp projections were compared regarding: (1) concordance in predicting a climatically suitable area, and (2) agreement in predicting a climatically unsuitable area.

Analysing phylogenetic, environmental and geographical distances

We calculated phylogenetic distances among clades using the 'distTips' command in the 'adephylo' package. Environmental distance was calculated as the Euclidean distance between the clades along the PCA axes. When calculating geographical distance, we considered the centre of gravity of the genetic-clade distributions as their representative position. We based the measure on a least-cost path analysis, considering Mediterranean water resistance as 100 times higher than that of land, to avoid easy dispersal across the Mediterranean Sea. Values in distance matrices are not independent, so we used Mantel regression as it corrects partial matrix value dependence. We used the 'multi.mantel' command in the 'phytools' package to explain phylogenetic distance among clades by means of environmental and geographical distances, and we set the number of permutations to $n = 9999$. Mantel correlations with 9999 permutations were run with the 'mantel.rtest' command in the 'ade4' package. We used variation partitioning (Borcard *et al.*, 1992; Meier *et al.*, 2010) to single out the individual contributions of environmental and geographical distances in explaining phylogenetic relatedness.

RESULTS

Genetic clusters

The phylogenetic reconstruction of the eight genetically defined maritime pine clades based on SNPs (Fig. 1) mainly matched the one obtained from chloroplast markers (Bucci *et al.*, 2007). It grouped the eastern (G5), central (G6) and southern (G7) Spanish populations, while Atlantic Iberian Peninsula (G1), Atlantic France (G3) and Tunisia (G8) were clustered together. These two clusters are connected to the eastern populations (south-eastern France, Italy and Corsica) (G2). Finally, the Moroccan populations (G4) are the most genetically different in relation to all other clades.

These results indicate short genetic distances among G5/G6/G7, and among G1/G3/G8, while G2 is genetically more distant from them and G4 is the most genetically distant clade.

However, Tunisian populations (G8) present atypical behaviour, grouping with western European clades (G1 and G3) when we expected them to group with the eastern one (G2), based on the mitochondrial and chloroplast markers for which distinct sets of individuals were analysed (see Burban & Petit, 2003; Bucci *et al.*, 2007). We suspect a possible mistake in labelling the individuals from Tunisia (G8) used for SNPs and nuclear SSRs (nSSRs) (the same sample may have been used for both markers; see Jaramillo-Correa *et al.*, 2015).

Species distribution models

The models calibrated individually for each genetic clade achieved a better performance than the models calibrated at the species level. The five statistical algorithms performed well (all AUC scores above 0.89; see Table S2 in Appendix S2), although RF outcompeted the rest in all cases except for G8, for which GLM and GAM performed slightly better.

Overall, the ensemble models outperformed most individual algorithm models except for RF. Even with that exception, the relationship between maritime pine distribution and its relevant climatic variables is well captured by the ensemble models, which represent a sound method for including the variability originating from different statistical algorithms (Araújo & New, 2007; Grenouillet *et al.*, 2011).

More details on model performance can be found in Appendix S2. All projections from 45 individual algorithms and 9 ensemble models are presented in Appendix S4.

Niche overlap

Overall, the overlap values obtained with both geographical and environmental analyses were low, with the exception of a few clades (Table 1). In the area catalogued as suitable territory for maritime pine (681,911 km^2), 79.8% was assigned to only one specific genetic clade, indicating a generally limited overlap among clades. Consistently, only 15.9% of the area was suitable for two different genetic clades and 4.3% for three different genetic clades (see Fig. 2); no area was suitable for more than three clades. Most of the territory suitable for three different genetic clades was concentrated in Spain (east and central plateau), mainly corresponding to range areas suitable for clades G5/G6/G7 with a very small contribution for G2/G6/G7. There was also geographical overlap for the three clades G1/G4/G8 in northern Africa.

Table 1 Geographical and environmental overlap between genetic clades (G1–G8).

		Geographical overlap (%)							
Env. overlap (D metric)		G1	G2	G3	G4	G5	G6	G7	G8
	G1		6.35	0	1.61	0	0	0.28	3.02
	G2	0.02		0	0	0	0.11	0.32	0
	G3	0.1	0.01		0	0	0	0	0
	G4	0.3	0.03	0.06		0	0.04	0	1.86
	G5	0	0.03	0	0		41.85	22.25	0
	G6	0.01	0.04	0	0.01	0.3		17.64	0
	G7	0.01	0.02	0	0.01	0.41	0.24		0
	G8	0.32	0.04	0.11	0.18	0.02	0.03	0.04	

Geographical overlap (above the diagonal) is calculated as the percentage of the total area suitable for any of the two components of the pair that is suitable for both members at the same time. Environmental overlap (below the diagonal) is estimated by the *D* metric (Schoener, 1970). Values above 0 indicate overlap in geographical or environmental (bioclimatic) space among any pair of the eight clades.

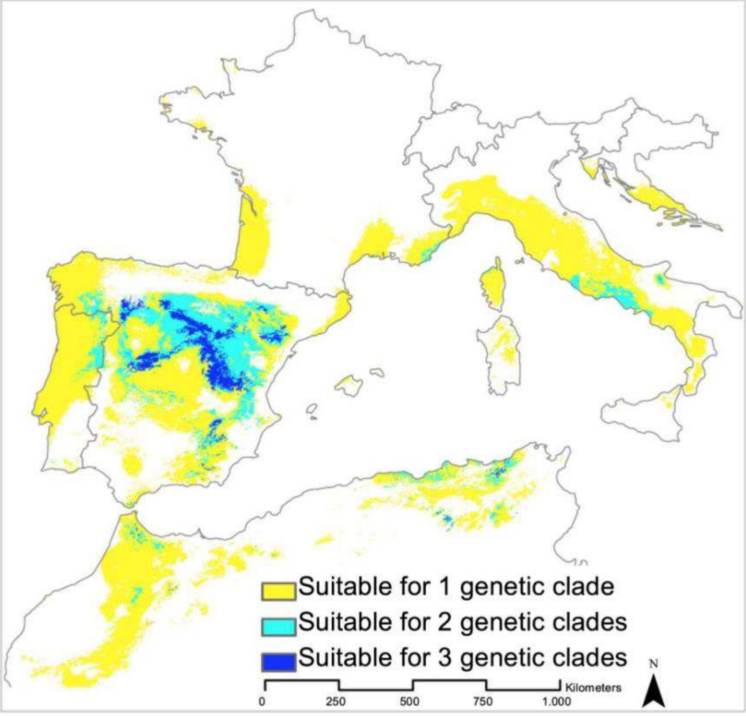


Figure 2 Geographical overlap of the projections for the different genetic clades (G1–G8). The suitability value indicates the number of genetic clades for which the area is climatically suitable. No territory is suitable for more than three genetic clades.

Clades G5, G6 and G7 located in eastern, central and southern Spain, respectively, showed particularly high niche overlap in both geographical and environmental space compared with the results obtained for the other genetic clades. We also found some overlap in environmental space among G1, G4 and G8, located on the Atlantic side of the Iberian Peninsula, Tunisia and Morocco, respectively. However, in these latter cases the overlap in environmental space did not translate into significant geographical overlap, suggesting that these environmental conditions appear in the study area only rarely.

The first two PCA-env axes explained 86.54% of the total variance (PC1 = 69.12% and PC2 = 17.24%; Fig. 3a), with PC1 being strongly correlated with summer water availability. Niche position and results (Fig. 3b) showed four distinguishable groups along the first two axes: G1/G3/G8, G5/G6/G7, G2 and G4. All equivalency tests were significant (P-value < 0.05), meaning that we could reject the null hypothesis that any pair of clade niches was equivalent. In addition, most similarity tests were not significant (except for G1/G8), so it cannot be assumed that the environmental conditions occupied by the

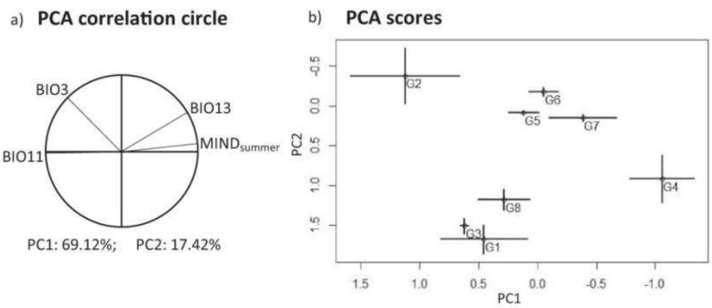


Figure 3 Ordination of climate space and projection of eight genetic clades. (a) Principal component ordination of the bioclimatic variables used in the study (BIO3, isothermality; BIO11, mean temperature of the coldest quarter; BIO13, precipitation of wettest month; and MINDsummer, moisture index in summer). (b) Niche position and breadth within the environmental PCA axes. Niche position is calculated as the median PCA score obtained for the occurrences of each genetic clade (G1–G8). Niche breadth is calculated as the variance among the occurrences of each genetic clade (G1–G8) considered for PC1 and PC2.

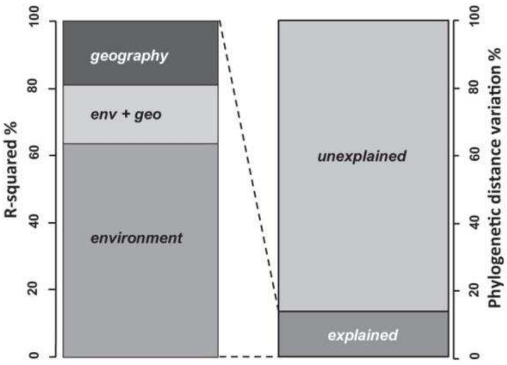


Figure 4 Variation partitioning of the contribution of the environment (from niche overlap) and geographical distance (from a least-cost path analyses between group geographical centres) explaining phylogenetic distances using a Mantel regression with $n = 9999$ permutations. The percentages address proportions explained from the full model containing both environmental and geographical distance predictors.

genetic clades were more similar than those expected by chance.

Mantel regressions on environmental and geographical distance matrices explained 12.6% of the variation in phylogenetic distance among the eight clades (Fig. 4). In them, the linear term for environmental distance was barely significant ($P = 0.06$), while the intercept ($P = 0.87$) and the geographical term ($P = 0.74$) were not (see the scatterplots of phylogenetic versus genetic distances and phylogenetic versus geographical distances in Appendix S2). Environmental distance explained *c.* 60% of the phylogenetic distance, while the contribution of geographical distance and the joint geography–environment distance contribution explained *c.* 20% each (Fig. 4). This was also reflected in higher and significant Mantel correlations between phylogenetic and environmental distances ($r = 0.32$, $P = 0.01$), as compared to geographical distances ($r = 0.21$, $P = 0.25$).

Constructing SDMs for species and molecular-based clades

In the comparison between SDMcm and SDMsp we detected a considerable mismatch between the two types of projections encompassing the species as a whole (Fig. 5). Notably, SDMcm projected habitat suitability for at least one of the molecular-based clades outside the projected geographical range covered by the species model. These areas contain conditions that are modelled as unsuitable from the species-level perspective, yet at least one of the clades found these regions suitable. Only very small areas predicted by the species model were not included within the niche of any of the eight clades.

DISCUSSION

SDMs, climatic niches and phylogeography

Both the strong niche differentiation and the low values of niche overlap among clades indicate niche evolution at the infraspecific level for maritime pine. Our results are in line with observed morphological differences among populations. They also support the subdivision of maritime pine into two subspecies (subsp. *atlantica* present in Atlantic plains, and the circum-Mediterranean subsp. *mesogeensis*) and several geographical varieties: var. *renoui* in Algeria and Tunisia, var. *maghrebiana* in Morocco, var. *corteensis* in Corsica, var. *hamiltonii* in the Mediterranean and var. *atlantica* in the Atlantic region (Barbéro *et al.*, 1998).

Furthermore, our study reinforces the conclusions of previous ones, in that both the phylogeographical pattern of maritime pine and the between-clade relationships largely relate to the demographic and migration history of the species, as most of the variation in phylogenetic distance cannot be explained by environmental or geographical distance. The existence of several

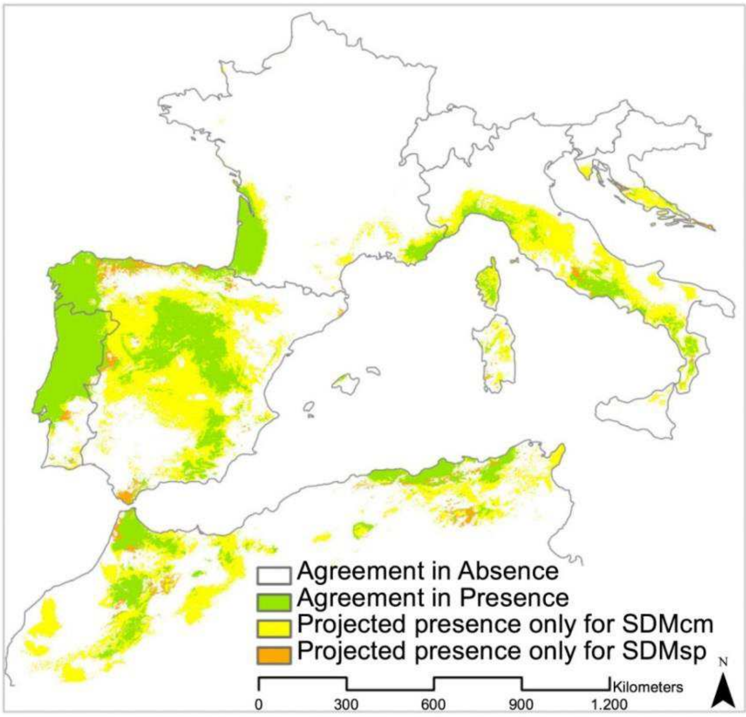


Figure 5 Comparison between projections of the species-level model (SDMsp not based on genetic clades) and the composite model (SDMcm based on molecular data). We distinguish four possible areas: agreement in presence, agreement in absence, projected presence only for SDMcm and projected presence only for SDMsp.

glacial refugia (in Africa, south-eastern Spain and the Atlantic coast of Portugal), together with their historical isolation, would be responsible for the delineated geographical pattern of the distinct genetic pools (Burban & Petit, 2003; Bucci *et al.*, 2007; Naydenov *et al.*, 2014).

Additional studies point to further evolutionary forces having played a role in shaping the regional structure of maritime pine, in particular selection. González-Martínez *et al.* (2002) detected higher differentiation in quantitative traits (i.e. probably driven by environmental factors) compared with molecular markers (i.e. a priori neutral) in populations from the western-most range of the species, while Jaramillo-Correa *et al.* (2015) detected a link between the genetic variation of potentially adaptive SNPs and the environmental heterogeneity at the regional scale. These results suggest that maritime pine populations have adapted to the novel and highly diverse ecological conditions after glaciation (see Cacho *et al.*, 2010, for a description of climatic deglaciation processes), and ultimately to different climatic niches across its current distribution.

The role of adaptation is further confirmed by our results, as the clusters defined in the environmental space (G1/G3/G8, G5/G6/G7, G2 and G4; Fig. 3) match the four basic clusters detected with the phylogenetic analyses (Fig. 1), pointing to an association between the environment and the delineation of the genetic clades. Additionally, environmental distance was shown to contribute more than geographical distance to explaining the variation of phylogeographical distance, thus suggesting that isolation by environment is more important than isolation by geographical distance in explaining population divergence in maritime pine. Finally, most of the area suitable for maritime pine (83.6%) has only been projected as suitable for one specific genetic clade, suggesting that each clade has specialized to the specific climatic conditions it occupies (as in D'Amen *et al.*, 2013, for African mammals). To sum up, these results indicate that regional environmental adaptation has contributed to the process of genetic diversification in maritime pine.

Evaluating the clusters in more detail reveals that the clades that are phylogenetically most closely related (G5/G6/G7 and G1/G8; see Fig. 1) show a certain degree of niche overlap, as indicated by their high estimates for environmental overlap (Table 1). Indeed, clades G5/G6/G7 represent a genetic admixture, as shown by the estimated membership coefficients (Q) from the structure clustering analysis (Appendix S1). The lack of major geographical and climatic barriers that would impede gene flow among these populations is reflected by the short genetic distances among them. A similar hypothesis can be drawn for clades G1/G3/G8, although it would be necessary to repeat the sampling for SNPs and nSSR for G8 (see Results). In their case, however, there are some geographical and climatic barriers that may have reduced gene flow and may thus have resulted in larger genetic distances. In this work, we have estimated environmental distances only within a bioclimatic framework. In future studies, it would be informative to include other factors relevant to determining species distribution if these results are to be confirmed. Edaphic properties, for example, are of great importance when studying the patterns of distribution of vegetation, as they directly influence resource availability. In our study, the lack of edaphic maps capturing the local variability of the soil inhabited by maritime pine prevented us from integrating such valuable information into SDMs.

Despite the general insight gained of specialized niches for the main genetic clades, it is also possible that all the clades share the same fundamental niche, i.e. the abiotic conditions in which a species can survive (Hutchinson, 1957). Dispersal limitations and local biotic interactions may have prevented them from occupying all available environments, which would have translated into different realized niches, i.e. the abiotic and biotic conditions in which a species can survive. The integration of information from common garden experiments with maritime pine (Alia *et al.*, 1997; Correia *et al.*, 2010) could help to further test whether the different genetic clades share the same fundamental niche, or if each of them has maintained differences in its fundamental niche behaviour.

Importance of molecular information in SDMs

Maritime pine occupies an extremely variable range of habitats, (climates types from Atlantic to Mediterranean, and elevations from sea level to 2000 m), and its distribution is restricted by multiple factors (e.g. wildfires, land-use policies prioritizing plantations and farming activities, overexploitation). Considering these constraints, and the fact that this study focused solely on climatic variables, we expected the SDMs to predict a suitable habitat extending beyond the observed distribution of the species. However, SDMsp only predicted a climatically suitable area for the species that was very similar to its actual range, and only when genetic information was included in the models (SDMcm) was the predicted spatial extent of suitable habitat largely broadened (Fig. 5). Because maritime pine has a broad climatic niche, and because our models do not constrain predictions geographically, the SDMcm, based on the SDMs for infraspecific clades, seems more realistic than the SDMsp projection. Similar results were obtained by Benito-Garzón *et al.* (2011), who found that including local adaptation and plasticity in SDMs also enlarged the area suitable for maritime pine in future climate scenarios when compared with classical SDMs.

Potential uses in the management of genetic resources

SDMs relate environmental variables to species distribution, thus providing a key tool for supporting conservation management decisions (e.g. to identify critical habitats, assess the probable future distribution of the species or select reserve zones). Integrating molecular information into SDMs is essential, as genetic diversity is a key component of biodiversity (Convention on Biological Diversity, http://www.cbd.int/convention/text). Our approach, integrating all available genetic species information into SDMs and assessing the relationship among genetic and ecological clusters, can benefit conservation and sustainable use of forest genetic diversity in at least in three essential aspects.

1. EUFORGEN has implemented a network of genetic conservation units based on ecogeographical zonation and the distribution of tree species in a given country (Koskela *et al.*, 2013). Infraspecific SDMs integrating molecular information can improve the definition of conservation units, as they provide spatially explicit information based on the performance of the species and its ability to colonize different ecological niches. In general, such models can help to define general conservation guidelines and to detect critically endangered species, genetic clades or areas (e.g. Thuiller *et al.*, 2011).
2. The use of genetic diversity in assisted migration programmes needs to be have a solid scientific basis. This approach highlights areas that are bioclimatically suitable for the species but not yet occupied. Thus, this method can help to select the territory and the appropriate genetic clade for extending the distribution of the species, minimizing the risk of failure.
3. Finally, this study contributes to our knowledge about regions of provenance (the basic unit for commercialization of seeds and plants in the European Union), currently already in implementation in many countries (see Alía *et al.*, 1996, for an example with maritime pine based on ecological, phenotypic and genetic similarities).

CONCLUSIONS

Climate change is nowadays guiding the conservation agenda, by helping to establish guidelines that ensure species survival. Forest trees, with their widespread distribution, are already facing a broad range of environmental conditions, which, as in the case of maritime pine, has led to distinct clades that are adapted to specific local climates. Understanding the ecological processes that have shaped the actual spatial structure of genetic clades is of great interest, not only for a better understanding of their adaptation to their native local environment but also to help to predict future behaviour and define conservation guidelines.

Our approach suggests that niche evolution is a major force driving the adaptive evolution of maritime pine clades. The broad spatial genetic structure that is observed today was most likely generated during the species' colonization history, including ecological and evolutionary processes and their interactions, starting from Ice Age refugia. The concordance among the environmental and genetic clusters, and the fact that the environment explains phylogenetic structure better than geographical distance, point to adaptation as a driver in shaping spatial genetic patterns in maritime pine. One key point remains to be tackled: whether the integration of provenance trials in SDMs would confirm this conclusion, or would rather point to habitat fragmentation or historical isolation as drivers of spatial genetic patterns.

ACKNOWLEDGEMENTS

We are grateful to Ricardo Alía and Santiago C. González-Martínez for the insightful discussion about this paper, and to Achilleas Psomas for his collaboration during the analysis stage. Thanks are extended to Rose Ruiz Daniels for the English language revision. We thank the COST Action FP1202 for enabling the collaboration between the different groups involved, which has supported the development of this paper. We acknowledge funding from the Spanish National Research Plan (AdapCon CGL2011-30182-C02-01), as well as grant FP7 NovelTree Breeding from the European Commission. M.S.V. was supported by a FPU from the Spanish Ministry of Education, Culture and Sport; D.G. was supported by a Ramón y Cajal Fellowship from the Spanish Ministry of Science and Innovation. We also thank two anonymous referees and two editors for their valuable help in improving the paper. Finally, we thank P. C. Grant-Pérez from Grant & Co. Academic English for revising the final manuscript.

REFERENCES

Alía, R., Martín, S., De Miguel, J., Galera, R., Agúndez, D., Gordo, J., Salvador, L., Catalán, G. & Gil, L. (1996) *Las regiones de procedencia de* Pinus pinaster *Aiton*. DGCONA, Madrid.

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

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.

Barbéro, M., Loisel, R., Quézel, P., Richardson, D.M. & Romane, F. (1998) Pines of the Mediterranean Basin. *Ecology and biogeography of* Pinus (ed. by D.M. Richardson), pp. 153–170. Cambridge University Press, Cambridge.

Bedia, J., Busqué, J. & Gutiérrez, J.M. (2011) Predicting plant species distribution across an alpine rangeland in northern Spain. A comparison of probabilistic methods. *Applied Vegetation Science*, **14**, 415–432.

Benito-Garzón, M., Alia, R., Robson, T.M. & Zavala, M.A. (2011) Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography*, **20**, 766–778.

Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045–1055.

Breiman, L. (2001) Random forests. *Machine Learning*, **45**, 5–32.

Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. (1984) *Classification and regression trees*. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.-J., Randin, C.F., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481–497.

Bucci, G., González-Martínez, S.C., Le Provost, G., Plomion, C., Ribeiro, M.M., Sebastiani, F., Alia, R. & Vendramin, G.G. (2007) Range-wide phylogeography and gene zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. *Molecular Ecology*, **16**, 2137–2153.

Burban, C. & Petit, R.J. (2003) Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Molecular Ecology*, **12**, 1487–1495.

Cacho, I., Valero-Garcés, B. & González-Sampériz, P. (2010) Revisión de las reconstrucciones paleoclimáticas en la Península Ibérica desde el último periodo glacial. *Clima en España: pasado, presente y futuro* (ed. by F.F. Pérez and R. Boscolo), pp. 9–24. CLIVAR-España; Ministerio de Medio Ambiente Rural y Marino.

Correia, I., Alia, R., Yan, W., David, T., Aguiar, A. & Almeida, M.H. (2010) Genotype × environment interactions in *Pinus pinaster* at age 10 in a multi-environment trial in Portugal: a maximum likelihood approach. *Annals of Forest Science*, **67**, 612p1–612p9.

D'Amen, M., Zimmermann, N.E. & Pearman, P.B. (2013) Conservation of phylogeographic lineages under climate change. *Global Ecology and Biogeography*, **22**, 93–104.

Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

González-Martínez, S.C., Alia, R. & Gil, L. (2002) Population genetic structure in a Mediterranean pine (*Pinus pinaster* Ait.): a comparison of allozyme markers and quantitative traits. *Heredity*, **89**, 199–206.

Grenouillet, G., Buisson, L., Casajus, N. & Lek, S. (2011) Ensemble modelling of species distribution: the effects of geographical and environmental ranges. *Ecography*, **34**, 9–17.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Hastie, T.J. & Tibshirani, R. (1990) *Generalized additive models*. Chapman & Hall/CRC, Boca Raton, FL.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hutchinson, G.E. (1957) *A treatise on limnology*. Wiley & Sons, New York.

Jaramillo-Correa, J.P., Rodríguez-Quilón, I., Grivet, D., Lepoittevin, C., Sebastiani, F., Heuertz, M., Garnier-Géré, P., Alia, R., Plomion, C., Vendramin, G. & González-Martínez, S. (2015) Molecular proxies of climate maladaption in a long-lived tree. *Genetics*, **199**, 793–807.

Koskela, J., Lefèvre, F., Schueler, S., Kraigher, H., Olrik, D.C. & Hubert, J. (2013) Translating conservation genetics into management: pan-European minimum requirements for dynamic conservation units of forest tree genetic diversity. *Biological Conservation*, **157**, 39–49.

Köble, R. & Seufert, G. (2001) Novel maps for forest tree species in Europe. *Proceedings of the 8th European Symposium on the Physico-Chemical Behaviour of Air Pollutants: 'A changing atmosphere!'*, pp. 17–20. Torino, Italy. Available at: ftp://mars.jrc.ec.europa.eu/Afoludata/Public/DOCU66/description_tree_species_maps.pdf

McCullagh, P. & Nelder, J.A. (1989) *Generalized linear models*. Chapman and Hall, London.

Meier, E.S., Kienast, F., Pearman, P.B., Svenning, J.-C., Thuiller, W., Araújo, M.B., And, A.G. & Zimmermann, N.E. (2010) Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography*, **33**, 1038–1048.

Moritz, C. (1994) Defining evolutionarily-significant-units for conservation. *Trends in Ecology and Evolution*, **9**, 373–375.

Naydenov, K.D., Alexandrov, A., Matevski, V., Vasilevski, K., Naydenov, M.K., Gyuleva, V., Carcaillet, C., Wahid, N. & Kamary, S. (2014) Range-wide genetic structure of maritime pine predates the Last Glacial Maximum: evidence from nuclear DNA. *Hereditas*, **151**, 1–13.

Nei, M. (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Oney, B., Reineking, B., O'Neill, G.A. & Kreyling, J. (2013) Intraspecific variation buffers projected climate change impacts on *Pinus contorta*. *Ecology and Evolution*, **3**, 437–449.

O'Neill, G.A., Hamann, A. & Wang, T. (2008) Accounting for population variation improves estimates of the impact of climate change on species' growth and distribution. *Journal of Applied Ecology*, **45**, 1040–1049.

Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008) Niche dynamics in space and time. *Trends in Ecology and Evolution*, **23**, 149–158.

Pearman, P.B., D'Amen, M., Graham, C.H., Thuiller, W. & Zimmermann, N.E. (2010) Within-taxon niche structure: niche conservatism, divergence and predicted effects of climate change. *Ecography*, **33**, 990–1003.

Pease, K.M., Freedman, A.H., Pollinger, J.P., McCormack, J.E., Buermann, W., Rodzen, J., Banks, J., Meredith, E., Bleich, V.C., Schaefer, R.J., Jones, K. & Wayne, R.K. (2009) Landscape genetics of California mule deer (*Odocoileus hemionus*): the roles of ecological and historical factors in generating differentiation. *Molecular Ecology*, **18**, 1848–1862.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

R Development Core Team (2013) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Rodríguez-Sánchez, F., Hampe, A., Jordano, P. & Arroyo, J. (2010) Past tree range dynamics in the Iberian Peninsula inferred through phylogeography and palaeodistribution modelling: a review. *Review of Palaeobotany and Palynology*, **162**, 521.

Santos-del-Blanco, L., Climent, J., González-Martínez, S. & Pannell, J.R. (2012) Genetic differentiation for size at first reproduction through male versus female functions in the widespread Mediterranean tree *Pinus pinaster*. *Annals of Botany*, **110**, 1449–1460.

Schoener, T.W. (1970) Nonsynchronous spatial overlap of lizards in patchy habitats. *Ecology*, **51**, 408–418.

Sierra-de-Grado, R., Pando, V., Martínez-Zurimendi, P., Peñalvo, A., Báscones, E. & Moulia, B. (2008) Biomechanical differences in the stem straightening process among *Pinus pinaster* provenances. A new approach for early selection of stem straightness. *Tree Physiology*, **28**, 835–846.

Takezaki, N., Nei, M. & Tamura, K. (2010) POPTREE2: software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Molecular Biology and Evolution*, **27**, 747–752.

Thomassen, H.A., Cheviron, Z.A., Freedman, A.H., Harrigan, R.J., Wayne, R.K. & Smith, T.B. (2010) Spatial modelling and landscape-level approaches for visualizing intra-specific variation. *Molecular Ecology*, **19**, 3532–3548.

Thuiller, W., Albert, C., Araújo, M.B., Berry, P.M., Cabeza, M., Guisan, A., Hickler, T., Midgley, G.F., Paterson, J., Schurr, F.M., Sykes, M.T. & Zimmermann, N.E. (2008) Predicting global change impacts on plant species' distributions: future challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 137–152.

Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009) BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369–373.

Thuiller, W., Lavergne, S. & Roquet, C. (2011) Consequences of climate change on the tree of life in Europe. *Nature*, **470**, 531–534.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868–2883.

Warren, D.L., Cardillo, M., Rosauer, D.F. & Bolnick, D.I. (2014) Mistaking geography for biology: inferring processes from species distributions. *Trends in Ecology and Evolution*, **29**, 572–580.

Wiens, J.J. & Graham, C.H. (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, **36**, 519–539.

Zimmermann, N.E., Edwards, T.C., Moisen, G.G., Frescino, T.S. & Blackard, J.A. (2007) Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *Journal of Applied Ecology*, **44**, 1057–1067.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Appendix S1 Detailed description of genetic clades

Table S1 Estimated membership coefficients (Q) for each individual within each cluster.

Appendix S2 Detailed methods and specifications for the preparation of presence/pseudo-absence data inputs, algorithm settings for the species distribution models, model performance and distance scatterplots.

Figure S1 Combination of EUFORGEN and Tree Species Distribution for Europe data bases for *Pinus pinaster* distribution.

Figure S2 Scatterplots of phylogenetic versus environmental distances and phylogenetic versus geographical distances.

Table S2 Assessment of model performance by means of the true skill statistic, area under the curve and H-measure scores.

Table S3 True skill statistic, H-measure, cutoff, sensitivity and specificity scores of the ensemble model (average of the five algorithms), for the species level model and for the models for the different genetic clades (G1–G8).

Appendix S3 Detailed description of bioclimatic variables and the selection procedure

Appendix S4 Geographical projections of the five algorithms (generalized linear model, general additive model, random forest, classification tree analysis and MaxEnt) and the ensemble model of maritime pine and its genetic clades (G1–G8).

Figure S3–S11 Geographical projection of *P. pinaster* and its eight genetically defined clades using five algorithms.

BIOSKETCH

This study represents a collaborative effort aimed at integrating genetic information in ecological niche modelling within the framework of the ADAPCON project. Within this project, we investigate patterns of adaptive variation in mediterranean conifers, integrating genetic variation (neutral and potentially adaptive) in natural populations, phenotypic variation in common garden experiments and ecological niche modelling. This work will be included in the first author's PhD thesis.

Editor: Linda Beaumont

Chapter 2

Assessing vulnerability of two iconic Mediterranean conifers to support European genetic conservation management in the face of climate change.

Assessing vulnerability of two iconic Mediterranean conifers to support European genetic conservation management in the face of climate change.

Serra-Varela *et al.*, in revision in Diversity and Distributions

Running head: Vulnerability as a tool for conservation

Serra-Varela, M.J.[1,2,3] maje.serra@gmail.com

Alía, R.[2,3] alia@inia.es

Ruiz Daniels, R.[3] rose_ruizdaniels@fastmail.fm

Zimmermann, N.E.[4,5] niklaus.zimmermann@wsl.ch

Gonzalo-Jiménez, J.[1,2*] jgonzalo@pvs.uva.es

Grivet, D.[2,3*] dgrivet@inia.es

*These authors contributed equally to the supervision of the work

1 – University of Valladolid, Department of Plant Production and Forest Resources, Avda. de Madrid 44, 34004 Palencia, Spain

2 - Sustainable Forest Management Research Institute, INIA - University of Valladolid, Avda. Madrid 44, 34004 Palencia, Spain.

3 – INIA, Forest Research Centre, Department of Forest Ecology and Genetics, Carretera A Coruña km 7.5, 28040 Madrid, Spain.

4 – Landscape Dynamics, Swiss Federal Research Institute WSL, Zuercherstrasse 111, CH-8903 Birmensdorf, Switzerland.

5 – Department of Environmental Systems Science, Swiss Federal Institute of Technology ETH, CH-8092 Zürich, Switzerland.

Corresponding author: Gonzalo-Jiménez, J. jgonzalo@pvs.uva.es; University of Valladolid, Department of Plant Production and Forest Resources, Avda. de Madrid 44, 34004 Palencia, Spain; Phone: 979 10 84 16; Fax: 979 10 84 40

Abstract

Aim: To assess two major components of vulnerability to climate change: adaptive capacity (approached by genetic clades) and exposure (approached by risk of habitat loss) illustrated by the maritime (*Pinus pinaster* Ait.) and Aleppo (*Pinus halepensis* Mill.) pines.To integrate such information in the selection of conservation strategies (*i.e. ex situ* vs *in situ* conservation) and to evaluate current European conservation efforts of forest genetic resources.

Location: Mediterranean Basin and European Atlantic coast.

Methods: We consider genetically defined clades as main conservation units. For each of them, we fit a species distribution model and project it to current climate and 42 different future climatic predictions for 2050. We create future suitability maps to assess risk of habitat loss based on the number of future climate projections for which each clade is suitable. According to this assessment on the risk of habitat loss, we propose suitable conservation strategies selecting amongst: *in situ*, *in situ* with monitoring and *ex situ* conservation.

Results: We found areas suitable for i*n situ* conservation for most of the clades, the exception being the central-eastern-southern clades of maritime pine and the Moroccan clade of Aleppo pine which required *ex situ* conservation. In the current European conservation network, three and two clades for maritime and Aleppo pine respectively, remain unrepresented, and the representation of the rest is unbalanced.

Main conclusions: We provide a tool to support conservation management of forest trees, an increasingly important task given the negative impact of climate change on forest ecosystems. We also provide a framework to increase the efficiency of the European conservation network: (i) exposure assessment should be considered as a requirement for a population to become a dynamic conservation unit (DCU); and (ii) as illustrated with for our two target species, the selection of DCUs should adequately represent all existing clades.

Keywords: Aleppo pine, Conservation biology, maritime pine*, Pinus halepensis* Mill., *Pinus pinaster* Ait., Species Distribution Models.

1. Introduction

Climate change has a worldwide impact on forest ecosystems often leading to their decline (*e.g.* Wang *et al.,* 2006; Allen *et al.,* 2010) and consequently to negative impacts on forest economies in many regions (Hanewinkel *et al.*, 2012). Therefore, and given the importance of preserving biodiversity (see http://www.cbd.int/convention/text), conservation plans are increasingly being implemented in national and international policies. In this context, it is essential to assess the extent to which a species or population is threatened by climate change (*i.e.* its vulnerability *sensu* Dawson *et al.,* 2011; see Mazziotta *et al.,* 2015 for an example). Nevertheless, the three components of vulnerability (adaptive capacity: ability to cope with climate change by persisting *in situ;* exposure: magnitude of climate change; and sensitivity: likelihood of an adverse response to climate change) are rarely considered together in conservation plans (Watson *et al.*, 2013), mainly due to the challenge in compiling the necessary information and the difficulty derived from combining various approaches. *Quantifying vulnerability is challenging,* especially in long-lived organisms such as trees where evaluating the viability of a population from demographic analyses, or estimating adaptive capacity by direct experimental observations is problematic.

Adaptive capacity of a species is determined by different factors (see Nicotra *et al.*, 2015 for a review), among which phenotypic plasticity and standing genetic variation play a fundamental role (Chevin *et al.*, 2010). In the particular case of forest trees, genetic variation is a key component of adaptive capacity (Hampe & Petit, 2005; Savolainen *et al.*, 2007; Neale & Kremer, 2011), and can be used as a proxy to identify genetic units (or clades) within which populations have a similar adaptive potential (see Serra-Varela *et al.,* 2015 for an example with *Pinus pinaster Ait.*). These units are relevant from a conservation perspective as they ideally represent the complete evolutionary potential and heritage of a species (Moritz, 1994) ultimately constituting conservation units (CUs).

Exposure has been addressed by different approaches (see Johnston *et al.,* 2009; Coops & Waring, 2011; Schueler *et al.,* 2014 for examples). In particular, Schueler *et al.*, (2014) analysed the exposure component of vulnerability on dynamic conservation units (DCUs) of the EUFORGEN programme (Koskela *et al.*, 2013) of six forest tree species, by calculating the increment between current and future favourability by means of species distribution models (SDMs). *I*n addition, they approached the exposure of the whole Forest Genetic European conservation network by estimating the velocity of climate change as proposed in Loarie *et al.*, (2009). However, their approach did not explicitly account for the adaptive component of vulnerability (*i.e.* infraspecific genetic variation), which is ultimately relevant as infraspecific clades have shown to differ in their response to climate change (*e.g.* Wang *et al.*, 2006; Benito-Garzón *et al.,* 2011; D'Amen *et al.,* 2013*).* Furthermore, they did not deal with future climate uncertainties derived from the wide range of future climate predictions originating from climate models and scenarios. Indeed, previous approaches that have attempted to integrate future predictions to assess best suitable provenances for the future or optimal conservation management options, generally fail to deal with the issue of uncertainty in future climate predictions (*e.g.* Wang *et al.,* 2006; Schueler *et al.,* 2013; Rehfeldt *et al.,* 2015). In this framework, SDMs (see Guisan & Zimmermann, 2000) constitute a suitable tool as they assess a species' habitat suitability under different climatic scenarios. However, the fact that SDMs generally do not account for local adaptation or migration processes (but see Normand *et al.,* 2011; Meier *et al.,* 2012; Hamann & Aitken, 2013), must be taken into consideration when interpreting the results. Still, SDMs enable to assess the risk of loss of suitable habitat (as a proxy of exposure) as they can predict whether future climatic conditions would be suitable or not for a species at any location in space.

Finally, considering sensitivity involves assessing the influence of climatic changes on the survival, persistence, fitness, performance or regeneration of a species or population (Dawson *et al.*, 2011). These issues are largely related to ecophysiology and thus, mechanistic models are suitable tools to address them. Nevertheless, these approaches require large empirical datasets and are difficult to implement in order to address vulnerability at a coarse scale.

In this study, we estimate vulnerability focusing on two of its components, adaptive capacity and exposure. We centre our approach on forest trees and aim to provide a tool to support conservation management decisions directed at maintaining species' evolutionary potential, and thus at increasing their probabilities to cope with climate change. We use exposure of a given genetic clade (a proxy of conservation units based on molecular markers) to select the most adequate among the different available conservation strategies (*i.e. in situ vs ex situ conservation; see* Ledig, 1986 for more detailed information) while dealing with future climate uncertainties (see Fig. 1). An optimized selection of a conservation strategy is ultimately relevant as DCUs aim to maintain evolutionary processes within tree populations, in order to safeguard their potential for future adaptation and keep them as dynamic evolving populations, both *in situ* and *ex situ*. In the face of climate change, the selection of conservation strategy over another should rely on a solid assessment of exposure, with the aim of maximizing the likelihood of effectively preserving a species. *We* deal with adaptive capacity by using genetically homogeneous groups of populations as conservation units. On the other hand, exposure is assessed by risk of habitat loss due to climate change by combining a wide range of future climate predictions (42) to incorporate the uncertainty in future climate predictions.

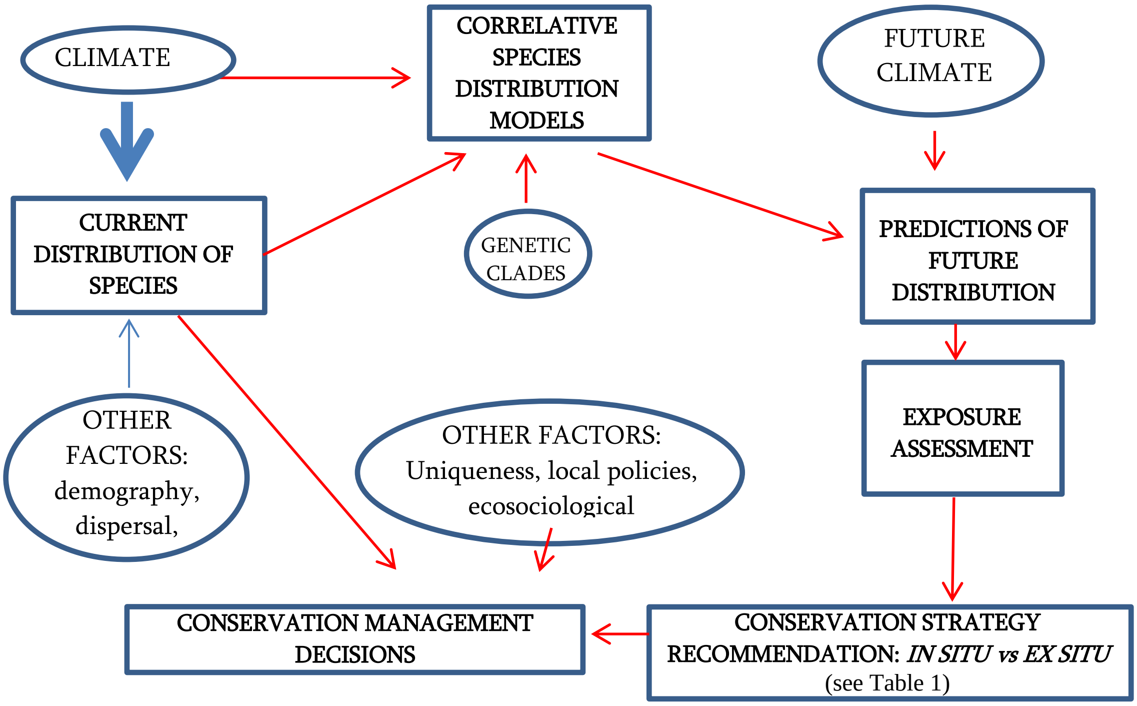


Figure 1: Framework for conservation management

To illustrate the performance of our proposed approach we select two Mediterranean pine species with very different evolutionary histories and genetic structure patterns, namely maritime pine (*Pinus pinaster* Ait. - see Bucci *et al.,* 2007; Jaramillo-Correa *et al.,* 2015) and Aleppo pine (*Pinus halepensis* Mill. - see Morgante *et al.,* 1998; Gómez *et al.,* 2001; Grivet *et al.,* 2009). As conservation units (CUs) we use clades defined with molecular markers for both species. The choice of these two species, fulfils two objectives: (i) it increases the representation of Mediterranean species, scarcely contemplated in current conservation approaches (Lefèvre *et al.*, 2013), and (ii) it enables the evaluation of current efforts in conserving forest tree genetic resources in Europe for these two species, by assessing the exposure of their currently defined DCUs to future climate change and whether all CUs (*i.e.* genetic clades) are appropriately represented.

2. Methods

2.1. Molecular data and definition of genetic clades

We obtained eight genetically defined clades for the full distribution range of *P. pinaster* from Serra-Varela *et al.* (2015) (Fig. 2a) namely Atlantic Iberian Peninsula (G1-pin), Eastern populations (G2-pin), Atlantic France (G3-pin), Morocco (G4-pin), Eastern (G5-pin), Central (G6-pin) and Southern (G7-pin) Spain, and Tunisia (G8-pin) based on mitochondrial, chloroplast and nuclear (Simple Sequence Repeats - SSRs -and Single Nucleotide Polymorphisms – SNPs) molecular markers. The SNP dataset comprised 772 individuals from 36 populations (see Jaramillo-Correa *et al.,* 2015 for more details).

For *P. halepensis* we detected seven different genetic clades namely Central and southern Spain clade (G1-hal), Balearic and southern France clade (G2-hal), Tunisian and northern Italian clade (G3-hal), Moroccan and southern Spain clade (G4-hal), Greek clade (G5-hal), Central and northern Spain clade (G6-hal) and Northern Spain and southern France clade (G7-hal) (see Fig. 2b), by performing a Bayesian clustering analysis using STRUCTURE (Pritchard *et al.*, 2000) on a SNP dataset (1325 individuals from 49 populations) covering most of the species natural range (see Appendix S1 in Supporting Information for more details).

Contrary to *P. pinaster* where the clades were spatially differentiated, *P. halepensis* presented transition areas occupied by more than one clade simultaneously and that we defined based on Q values as explained in Appendix S1. Some small areas of the distribution of the species (located in Algeria) could not be assigned to any clade due to insufficient sampling (see Fig. 2b).

For both, *P. pinaster* and *P. halepensis*, the clades were based on a *priori* neutral molecular markers, which are primarily influenced by demographic processes and not by adaptation. However, some of these markers may also be influenced by adaptive selection (especially the SNPs, *e.g.* Jaramillo-Correa *et al.,* 2015 for *P. pinaster).*

2.2.-Species data

The complete native range for both species was obtained by combining the Tree Species Distribution for Europe (TSDE; Köble & Seufert, 2001) from the Joint

Research Center's AFOLU data portal (ftp://mars.jrc.ec.europa.eu/Afoludata/Public/DS66/) and the EUFORGEN database from the European forest genetic resources programme (http://www.euforgen.org/distribution-maps/) (see Serra-Varela *et al.* 2015 for further details).

We prepared a presence-pseudoabsence dataset for each genetic clade individually. Presences of genetic clades were defined as the subset of the overall presences records that belonged to one specific genetic clade. In the case of *P. halepensis* presence records of transition zones were considered as presence records for both genetic clades inhabiting that territory. Possible pseudo-absences corresponded to all the rest of the territory within the study area where TSDE reported 0 % occupancy.

The number of presences for the genetic clades of *P. pinaster* amounted to: 59,159 (G1-pin), 12,225 (G2-pin), 17,898 (G3-pin), 4,966 (G4-pin), 8,792 (G5-pin), 13,005 (G6-pin), 4,896 (G7-pin), and 4,058 (G8-pin). For *P. halepensis* the number of presence records was as follows: 34,251 (G1-hal), 8,010 (G2-hal), 24,738 (G3-hal), 12,228 (G4-hal), 4,390 (G5-hal), 48,369 (G6-hal), 22,720 (G7-hal). The selection method and the number of selected pseudo-absences are specified below.

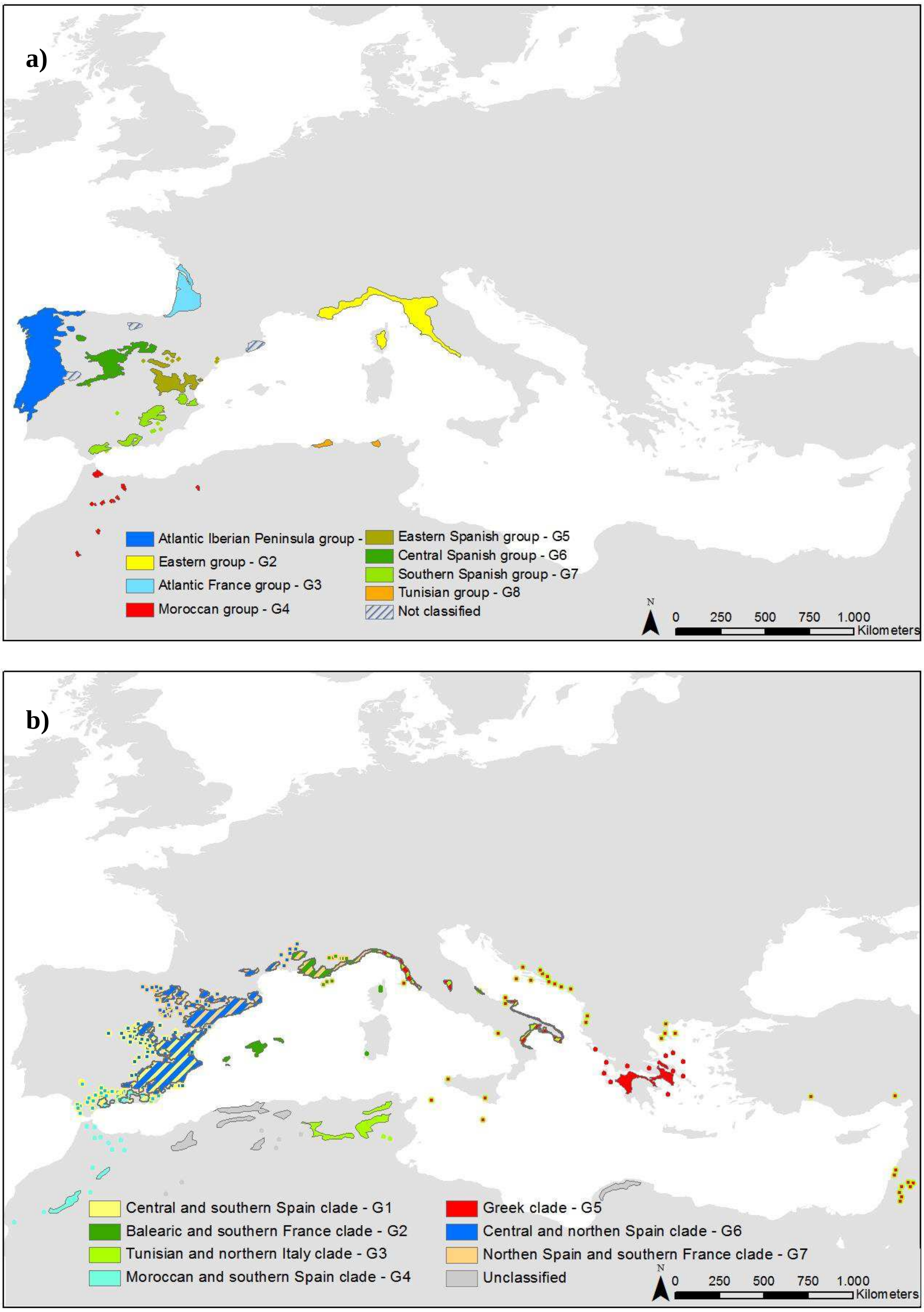


Figure 2: Distribution of the genetic clades of *Pinus pinaster* Ait. from Serra-Varela *et al.* (2015) (a) and *Pinus halepensis* Mill. (b) along the natural distribution of the species.

2.3.-Bioclimatic data

We used the 19 bioclimatic variables available in WORLDCLIM (Hijmans *et al.*, 2005) representative of the period between 1950-2000 for the analysis. After screening for correlations and variance inflation, the final set of relevant variables that correlate less than 0.75 among each other contained: BIO4 (Temperature Seasonality), BIO11 (Mean Temperature of Coldest Quarter), BIO12 (Annual Precipitation) and BIO18 (Precipitation of Warmest Quarter) (see Appendix S2 for more details). Overall, we detected a strong relationship between this same set of bioclimatic predictors and the distribution of each species.

Future bioclimatic predictions were also obtained from WORLDCLIM, as these predictions are based on the most recent Global Climate Models (GCMs) projections that have been used in the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment report (IPCC, 2013). We used all GCMs that were simultaneously available for the scenarios of representative concentration pathways (RCP) 2.6, 4.5 and 8.5. This included 14 GCMs (see Appendix S2) and totalled to 42 different future climate predictions.

2.4.-Species Distribution Models

We used five different statistical methods, namely General Linear Model (GLM - McCullagh & Nelder, 1989), Generalized Additive Model (GAM - Hastie & Tibshirani, 1990), Random Forest (RF - Breiman, 2001), Classification Tree Analysis (CTA - Breiman *et al.,* 1984) and MaxEnt (Phillips *et al.*, 2006) in an ensemble approach by means of the same settings and statistical environment as specified in Serra-Varela *et al.* (2015). Model performance was assessed by means of the True Skill Statistic (TSS; Allouche *et al.,* 2006). Thus, we built 8 ensemble models for *P. pinaster* clades (G1-pin to G8-pin, based on 5x8=40 individual algorithm models) and 7 ensemble models for *P. halepensis* clades (G1-hal to G7-hal, based on 5x7=35 individual algorithm models).

We set the number of randomly selected pseudo-absences for all the models to five times the number of total presences used per species (see Serra-Varela *et al.* 2015 for more details). This amounted to 644,190 presence and pseudo-absence points for *P. pinaster* and 679,110 for *P. halepensis*. Pseudo-absences and presences were given weights inversely proportional to their respective numbers, so as to give equal total weights to the two sets as recommended by Barbet-Massin *et al.,* (2012).

The large number of occurrence records available permitted a random division of each dataset (corresponding to both species and to each clade) into two equally-sized subsets for training and evaluation, and maintaining the prevalence.

The probabilistic model outputs based on current and future climatic conditions were converted to binary presence-absence maps by defining thresholds that maximized TSS values in a test of initial models against the half of the data not used for model building (split-sample test). Thereby for each climatic dataset (current plus 42 future climate predictions representing 2050) eight/seven different distribution maps were created respectively for *P. pinaster* and *P. halepensis* corresponding to each of the genetically-defined clade models.

2.5.-Future suitability maps

For each clade, we obtained 42 future binary projections, which classify each cell as environmentally suitable or unsuitable by 2050. None of the 42 future climate predictions is better than another (but see Fordham *et al.,* 2011). Instead, all are equally likely and we used the number of future climate predictions projecting suitability in one specific cell as an indicator of the degree of agreement among models that the future habitat will be suitable in that cell. Thereby, combining the 42 binary projections, we obtained a map in which values could possibly range from 0.0 (none of the future climate predictions was projected to be suitable) to 1.0 (all future climate predictions tested were projected to be suitable) with higher scores indicating higher agreement of suitable habitat in the future. Then, we defined three different future suitability categories: “likely suitable” with suitability scores >0.7 (suitable habitat for more than 70% of future projections), “uncertain” with a suitability score of 0.36-0.7, and “likely unsuitable” with suitability scores <0.36. We performed this analysis for each clade individually.

2.6.-Defining conservation strategies based on exposure to climate change

Future suitability maps were overlaid with maps of current distribution in order to assess risks of habitat loss (a proxy of exposure), and set the basis for the recommendation of the different conservation strategies within the current distribution of each clade (see Table 1). *In situ* conservation is an effective conservation strategy for populations highlighted as "likely suitable" and "uncertain" by future suitability maps. However, in this last case ("uncertain" locations), a monitoring programme (see Graudal *et al.,* 2014) would be essential in order: (i) to ascertain that genetic processes linked to the adaptability of the species are maintained over time, and (ii) to apply more intense management aimed at supporting local adaptation processes if necessary. On the contrary, *ex situ* conservation is recommended in the case of exposed populations (Schueler *et al.*, 2014) that are classified as "likely unsuitable". This strategy includes translocation (Leech *et al.*, 2011) and/or conservation in germplasm banks. Between both, translocation is the preferred option as it maintains dynamic evolution within populations and, future suitability maps identify possible locations lowly exposed in the future. Only when future suitability maps do not highlight available locations for translocation, germplasm banks are recommended.

Finally, for each clade, we calculated the percentage of the currently occupied territory proposed for *ex situ* conservation in order to assess the risk of habitat loss of the clade.

2.7.-Exposure assessment of Dynamic Conservation Units

We assessed the exposure of the currently defined DCUs for both species covering the current range of the species which totalled to 46 in the case of *P. pinaster* - 36 from the EUFGIS database and 10 from Rodríguez-Quilón *et al.,* (2016) based on molecular and quantitative trait information - and 19 for *P. halepensis* from the EUFGIS database. DCUs from EUFGIS without a clear association to a particular genetic clade were excluded from the analysis.

We associated each unit to its corresponding clade and assessed the degree to which different gene pools (clades) were represented in the European network. We also evaluated the risk of losing each DCU due to failure of finding suitable habitat in the future using future suitability maps. In the case of *P. halepensis*, when DCUs represented two clades simultaneously (transition zones) risk evaluations of habitat (and clade) loss were performed separately for both clades.

Table 1: Conservation strategy recommendations based on current distribution of the clade, current projection of its species distribution model (SDM) and future suitability map.

Current distribution	SDM - Current projection	Future Suitability Map 2050	Conservation Strategy
present	-	likely unsuitable	*ex situ*
present	-	likely suitable	*in situ*
present	-	uncertain	*in situ* with monitoring
absent	suitable	likely suitable	current first option area for translocation
absent	suitable	uncertain	current second option area for translocation
absent	unsuitable	likely suitable	midterm first option area for translocation
absent	unsuitable	uncertain	midterm second option area for translocation

3. Results

3.1.-Species Distribution Models

Models performed similarly to those described by Serra-Varela *et al.* (2015) in cross-validation tests: (i) models built by means of individual algorithms performed well (TSS and AUC values above 0.80 in all cases except for GLM G4-hal; see Table S2 and S3 in Appendix S2); (ii) RF displayed the highest AUC and TSS scores in general; (iii) overall, ensemble models obtained higher TSS scores than most single models from individual statistical algorithms, except for RF. With the obtained TSS values, the relationship between both species' distribution and their relevant climatic variables can be considered well captured by the ensemble models, which represent a sound method to include variability originating from different statistical algorithms (Araújo & New, 2007; Grenouillet *et al.,* 2011).

Finally, sensitivity and specificity scores obtained in the split-sample test of the final ensemble binary models were also very high (above 0.95 in all cases; see Tables S4 and S5 in Appendix S2). All geographical projections can be checked in Appendix S3. The projected niches were largely broader than the current distribution of the clades (see Appendix S3), especially in the case of *P. halepensis.*

3.2.-Future suitability maps

Likely suitable areas at clade level generally concentrated around the clade's current distribution (see Appendix S3). Away from the current distribution, uncertain habitat suitability areas were majorly found in other locations around the Mediterranean basin (*e.g.* G2-pin and G1-hal). Only in the case of G6-hal, we found larger regions of likely suitable areas in northern Central Europe. For both species, there were several other clades that additionally found suitable areas in northern Central Europe, although with a medium to low probability of suitability (which was classified as uncertain or likely unsuitable).

3.3.-Defining conservation strategies based on exposure to climate change

According to the results obtained, we proposed conservation guidelines for each genetic clade based on its characteristics with regards to their currently observed and their projected current and future distribution (see Appendix S3).

As a first remarkable result, the risk of habitat loss of the clades varied widely from one clade to another within each of the two species. We detected from very slightly exposed clades, in which less than 15 % of the currently realized niche was recommended for *ex situ* conservation (*e.g.* G8-pin and G3-hal; see Table S6 in Appendix S3), to cases in which over 90 % of the currently realized niche was recommended for *ex situ* conservation (*e.g.* G5/G6/G7-pin or G4-hal). This analysis revealed very diverging patterns of exposure to possible future habitat loss for both species along the Mediterranean coast of Spain: while *P. pinaster*'s clades inhabiting this area (G5/G6/G7-pin) were highly exposed, *P. halepensis'* clades occupying the same territory (G1/G4/G6/G7-hal) were not (percentage below 60 %, except for G4-hal in southern Spain and Morocco, which was also found to be highly exposed).

Overall, we found that *in situ* conservation should be suffice to guarantee conservation of the different clades, as large areas were found in which *in situ* conservation will likely be successful without need of special monitoring programmes. Nevertheless, there were some cases in which monitoring seemed strictly necessary (*i.e.* the cases of G3-pin, G4-pin and G5-hal).

There were few clades, which lacked areas that could be proposed for *in situ* conservation (*i.e.* G5/G6/G7-pin and G4-hal). In these cases it was necessary to consider *ex situ* conservation strategies either by translocation or by conservation in germplasm banks. For three of them (G5/G7-pin and G4-hal) translocation was the recommended *ex situ* conservation strategy, while for G6-pin that lacked translocation areas, germplasm bank conservation had to be recommended. In addition, because there were only few translocation areas available for G7-pin, the conservation of this clade would also benefit from germplasm conservation.

3.4.-Exposure assessment of Dynamic Conservation Units

First, we analysed whether the already established EUFORGEN DCUs for *P. pinaster* and *P. halepensis* adequately represented the different gene pools detected in both species. In the case of *P. halepensis*, DCUs located within transition zones were considered representative for both clades. We found that while some clades were overrepresented (*e.g.* in *P. pinaster* 13 and 10 out of 36 DCUs harbour G2/G7-pin populations respectively and in *P. halepensis* 14 out of 19 DCUs combined admixtures of G3/G5-hal populations), there were other clades not included in the

DCU network (G1/G4/G8-pin and G1/G4-hal). The remaining clades, although present within the DCU network, were underrepresented (*e.g.* G6-pin and G6-hal with just 2 and 1 DCUs, respectively; see Appendix S4 for more details).

Second, we evaluated the risk of habitat loss within the different DCUs. In the case of *P. pinaster*, we detected seven DCUs classified as "likely suitable" in 2050 all of which belonged to G2-pin (except for one representing G5-pin). The other DCUs were classified as "likely unsuitable" (13) or "uncertain" (16) (see Fig. 3 and Appendix S4 for more details). As for *P. halepensis*, we detected eight DCUs classified as "likely suitable" in 2050 for one of the clades they represented (all of them located in Italy and representing G2/G3-hal). The rest were classified as "likely unsuitable" for the clade(s) they were representing but for one DCU classified as "uncertain". None of the DCUs were classified as "likely suitable" for both represented clades at the same time (see Fig. 3 and Appendix S4 for more details). Thus, except for some DCUs representing G2/G3-hal, the clades would need to be revised for conservation planning (selection of new DCUs or applying a more intense management to accelerate the adaptation processes within populations).

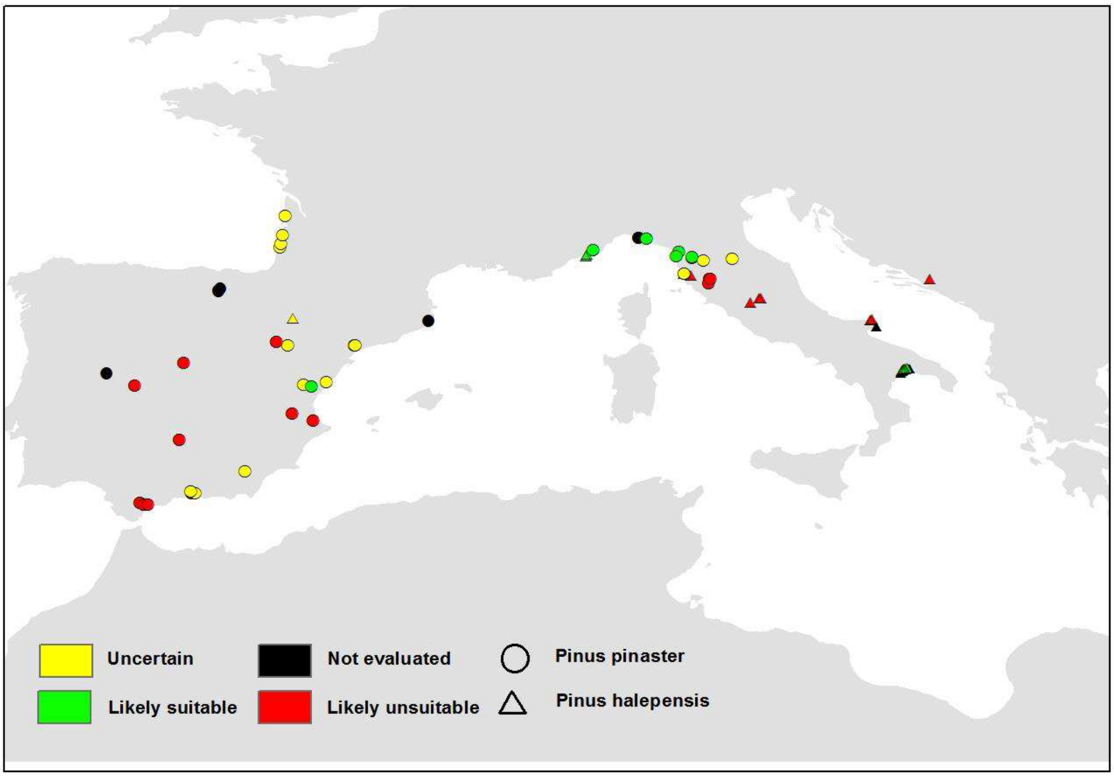


Figure 3 : Exposure assessment for the EUFGIS Dynamic Conservation Units (DCUs) of *Pinus pinaster* Ait. and *Pinus halepensis* Mill. along their natural distribution.

The ten genetically-homogeneous conservation relevant units for *P. pinaster*, were composed of an oscillating number of populations ranging from 1 to 10 (see Table S10 in Appendix S4). Overall, all genetics clades were represented except for G8-pin. Exposure varied widely across populations within units composed of more than one population (example SpAtl in which the three possible classifications where assigned along its ten composing populations) while among single population-groups the classification ranged from "likely unsuitable" to "uncertain".

4. Discussion

The consequences of climate change on European forest trees have been well documented (*e.g.* Thuiller *et al.,* 2006; Sykes *et al.,* 2009). Mediterranean regions in particular are expected to suffer more intensively from the effects of climate change, compared to other more northerly regions of Europe (Lindner *et al.*, 2010) due to their location at the rear edge of species' distribution (Hampe & Petit, 2005), and to their expected increased risk of drought and fire (Mouillot *et al.*, 2002; Pausas, 2004). Despite this threat, Mediterranean species remain underrepresented in the current European conservation network (Lefèvre *et al.*, 2013) as well as in earlier conservation studies (*e.g.* Schueler *et al.,* 2014).

To fill this gap, we employ two iconic Mediterranean species to illustrate an approach that assesses two major components of vulnerability to climate change (adaptive capacity and exposure), and that provides solid scientific criteria to support conservation management decisions. The goal of conservation management should not only be the preservation of species in climatically stable areas as proposed by Iwamura *et al.,* (2010), but also to guarantee the maintenance of the species` potential to dynamically evolve and adapt to new emerging climatic conditions. By selecting genetic clades as evolutionary units and by assessing exposure accounting for future uncertainties, our approach constitutes a step forward in comparison with previous studies such as those of Hamann *et al.,* (2004), Kapeller *et al.,* (2012) and Schueler *et al.,* (2014).

Regarding the first component of vulnerability, *i.e.* adaptive capacity, our two target species have high levels of adaptive differentiation within clades - reported in Rodríguez-Quilón *et al.* (2016) for the clades of *P. pinaster* and likely to be also present within the gene pools of *P. halepensis* as shown in Voltas *et al.,* (2008) - as well as show significant levels of phenotypic plasticity (see Corcuera *et al.* (2010, 2011) for *P. pinaster*, and Baquedano *et al.* (2008), Santos-del-Blanco *et al.* (2013) for *P. halepensis*). Regarding the second component of vulnerability, exposure, according to our results, with few exceptions, most of the clades defined based on molecular data span relatively large areas of the species distribution that are likely to remain suitable in the future. Thus, it is expected that these clades will be able to withstand climate change and *in situ* conservation management should be the basic strategy to apply. Nevertheless, there are some highly exposed clades (*e.g.* G5/G6/G7-pin and G4-hal) for which further work is needed to disentangle the relationship between adaptive genetic variability and resiliency towards climatic changes. In these cases it is necessary to assess whether high exposure translates into high vulnerability, or on the contrary adaptive capacity of the clades is enough to cope with climatic changes and thus the clades are not ultimately threatened.

Our results reveal suitable habitat space extending beyond the current distribution of the clades. Reasons for this are, for instance, dispersal limitations, historical barriers, land use, soil factors, or other factors unaccounted for in our model, which only considers climatic factors. We detect this especially in the case of *P. halepensis* (see Appendix S3), and it illustrates that such factors should also be analysed and considered when developing a conservation plan. In addition, in our approach we only assessed exposure in terms of the effects of climate change on abiotic factors (*i.e.* climatic variables). Biotic interactions may also be altered along with climatic changes resulting in new pests or competitors constraining the distribution of species in the future. Finally, from a temporal point of view, we estimated exposure by means of future projections for a 30-year period (from present to 2050), which seems realistic in long-lived forest trees with long rotation times (from 30 to almost 100 years) as there is likely less than one generation present during this time span. Thus, we can assume niche conservatism between the two time-slices analysed, despite the fact that niche evolution has been reported to influence *P. pinaster* at the infraspecific level(Serra-Varela *et al.*, 2015).

In the framework of the current pan-European conservation network (EUFORGEN and EUFGIS programme), Koskela *et al.,* (2013) established the minimum requirements for DCUs of forest tree genetic diversity, namely: (i) to designate genetic conservation areas (ii) to set up a basic management plan, and (iii) to identify one or more species as targets to conserve genetic diversity. Here, we suggest the inclusion of a new factor as a compulsory minimum requirement: the overlay of current and future habitat suitability. Accounting for exposure is essential in a conservation network as it provides insights in to the most appropriate management of DCUs. For instance, if we are dealing with a population for which climate will likely/uncertainly become unsuitable in the future, monitoring would become an indispensable tool to detect population decay, or to address possible management with the aim of accelerating adaptive processes. If DCUs are not capable of tracking climate change, all resources invested in their conservation management would become obsolete. Furthermore, the size of DCUs should be estimated taking in to account the velocity of climate change (Loarie *et al.*, 2009; Hamann *et al.*, 2015), which was assessed by Schueler *et al.*, (2014) for the whole network, and species-specific requirements to maintain viable populations. We also highlight possible improvements for our two model species *P. pinaster* and *P. halepensis*: (i) new DCUs are necessary to represent all genetic clades of a given species, and (ii) in the specific case of *P. halepensis*, in which there are territories occupied by two clades simultaneously, it would be more cost-effective to select adequate DCUs for both clades at the same time.

In this work, we aimed to integrate all previous efforts related to conservation of genetic resources and apply them to two Mediterranean iconic species, in order to enhance the design of an optimized conservation network. Within this framework, we also included in our analysis the ten conservation relevant units highlighted for *P. pinaster* (see Rodríguez-Quilón *et al.,* 2016). Within units composed of more than one population, exposure assessment can guide the selection of the most relevant target population for conservation. On the other hand, as explained for DCUs, exposure assessment indicates adequate management options to increase the likelihood of successfully maintaining target populations.

To conclude, forest management can benefit from our approach by bridging both conservation and active management. In the two case studies presented here, we are able to identify areas with different vulnerability levels, and therefore where different managing options can be established to enhance resilience. Further, recommendations concerning target areas for translocation can be used to assign afforestation needs that may have objectives other than the conservation of biodiversity (such as habitat restoration, wood production or protection against erosion).

Acknowledgements

We are grateful to Santiago C. González-Martínez for insightful discussions. Thanks are extended to COST Action FP1202 for enabling the collaboration between the different groups involved, which has supported the development of this paper. We acknowledge funding from the Spanish National Research Plan AdapCon (CGL2011-30182-C02-01), as well as from the TipTree (BiodivERsA-ERANET). M.J.S.V. was supported by a FPU from the Spanish Ministry of Education, Culture and Sport; D.G. was supported by a Ramón y Cajal fellowship from the Spanish Ministry of Science and Innovation; R.R.D. was supported by a FPI from the Spanish Ministry of Economy and Competiveness; N.E.Z. was supported by an SNSF Grant (No: 31003A_149508/1). We also thank the EUFGIS national focal points and the EUFORGEN national coordinators for their inputs to development and maintenance of the EUFGIS database.

Biosketch

This study represents a collaborative effort aiming at integrating genetic information in ecological niche modelling within the framework of the AdapCon project. Within this project, we investigate patterns of adaptive variation in Mediterranean conifers, integrating genetic variation (neutral and potentially adaptive) in natural populations, phenotypic variation in common garden experiments and ecological niche modelling. This work will be included in the first author´s PhD thesis.

References

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

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, **22**, 42–47.

Baquedano, F.J., Valladares, F. & Castillo, F.J. (2008) Phenotypic plasticity blurs ecotypic divergence in the response of *Quercus coccifera* and *Pinus halepensis* to water stress. *European Journal of Forest Research*, **127**, 495–506.

Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **3**, 327–338.

Benito-Garzón, M., Alía, R., Robson, T.M. & Zavala, M.A. (2011) Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography*, **20**, 766–778.

Breiman, L. (2001) Random forests. *Machine Learning*, **45**, 5–32.

Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. (1984) *Classification and Regression Trees*, Wadsworth & Brooks/Cole Advanced Books & Software, Monterey.

Bucci, G., González-Martínez, S.C., Le Provost, G., Plomion, C., Ribeiro, M.M., Sebastiani, F., Alía, R. & Vendramin, G.G. (2007) Range-wide phylogeography and gene zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. *Molecular Ecology*, **16**, 2137–2153.

Chevin, L.-M., Lande, R. & Mace, G.M. (2010) Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biology*, **8**, e1000357.

Coops, N.C. & Waring, R.H. (2011) Estimating the vulnerability of fifteen tree species under changing climate in Northwest North America. *Ecological Modelling*, **222**, 2119–2129.

Corcuera, L., Cochard, H., Gil-Pelegrin, E. & Notivol, E. (2011) Phenotypic plasticity in mesic populations of *Pinus pinaster* improves resistance to xylem embolism (P50) under severe drought. *Trees - Structure and Function*, **25**, 1033–1042.

Corcuera, L., Gil-Pelegrin, E. & Notivol, E. (2010) Phenotypic plasticity in *Pinus pinaster* δ13C: environment modulates genetic variation. *Annals of Forest Science*, **67**, 812–812.

D'Amen, M., Zimmermann, N.E. & Pearman, P.B. (2013) Conservation of phylogeographic lineages under climate change. *Global Ecology and Biogeography*, **22**, 93–104.

Dawson, T.P., Jackson, S.T., House, J.I., Prentice, I.C. & Mace, G.M. (2011) Beyond predictions: biodiversity conservation in a changing climate. *Science*, **332**, 53–58.

Fordham, D.A., Wigley, T.M.L. & Brook, B.W. (2011) Multi-model climate projections for biodiversity risk assessments. *Ecological Applications*, **21**, 3317–3331.

Gómez, A., Alía, R. & Bueno, M.A. (2001) Genetic diversity of Pinus halepensis Mill . populations detected by RAPD loci. *Annals of Forest Science*, **58**, 869–875.

Graudal, L., Aravanopoulos, F.A., Bennadji, Z., Changtragoon, S., Fady, B., Kjær, E.D., Loo, J., Ramamonjisoa, L. & Vendramin, G.G. (2014) Global to local genetic diversity indicators of evolutionary potential in tree species within and outside forests. *Forest Ecology and Management*, **333**, 35–51.

Grenouillet, G., Buisson, L., Casajus, N. & Lek, S. (2011) Ensemble modelling of species distribution: The effects of geographical and environmental ranges. *Ecography*, **34**, 9–17.

Grivet, D., Sebastiani, F., González-Martínez, S.C. & Vendramin, G.G. (2009) Patterns of polymorphism resulting from long-range colonization in the Mediterranean conifer Aleppo pine. *The New Phytologist*, **184**, 1016–28.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.

Hamann, A. & Aitken, S.N. (2013) Conservation planning under climate change: Accounting for adaptive potential and migration capacity in species distribution models. *Diversity and Distributions*, **19**, 268–280.

Hamann, A., Aitken, S.N. & Yanchuk, A.D. (2004) Cataloguing *in situ* protection of genetic resources for major commercial forest trees in British Columbia. *Forest Ecology and Management*, **197**, 295–305.

Hamann, A., Roberts, D.R., Barber, Q.E., Carroll, C. & Nielsen, S.E. (2015) Velocity of climate change algorithms for guiding conservation and management. *Global Change Biology*, **21**, 997–1004.

Hampe, A. & Petit, R.J. (2005) Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, **8**, 461–467.

Hanewinkel, M., Cullmann, D.A., Schelhaas, M.J., Nabuurs, G.J. & Zimmermann, N.E. (2012) Climate change may cause severe loss in the economic value of European forest land. *Nature Climate Change*, **3**, 203–207.

Hastie, T.J. & Tibshirani, R. (1990) *Generalized additive models*, Chapman & Hall/CRC, Boca Raton - London - New York - Washington D.C.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of*

Climatology, **25**, 1965–1978.

IPCC (2013) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, (ed. by T.F. Stocker), D. Qin), G.K. Plattner), M. Tignor), S.K. Allen), J. Boschung), A. Nauels), Y. Xia), V. Bex), and P.M. Midgley) Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,.

Iwamura, T., Wilson, K.A., Venter, O. & Possingham, H.P. (2010) A climatic stability approach to prioritizing global conservation investments. *PLoS ONE*, **5**.

Jaramillo-Correa, J.P., Rodríguez-Quilón, I., Grivet, D., Lepoittevin, C., Sebastiani, F., Heuertz, M., Garnier-Géré, P., Alía, R., Plomion, C., Vendramin, G. & González-Martínez, S. (2015) Molecular proxies of climate maladaption in a long-lived tree. *Genetics*, **199**, 1–15.

Johnston, M.H., Campagna, M., Gray, P.A., Kope, H.H., Loo, J.A., Ogden, A.E., O'Neill, G.A., Price, D.T., Williamson, T.B. onM. H., Campagna, M., Gray, P.A., Kope, H.H., Loo, J.A., Ogden, A.E., O'Neill, G.A., Price, D.T. & Williamson, T.B. (2009) *Vulnerability of Canada's tree species to climate change and management options for adaptation: an overview for policy makers and practitioners.*, Canadian Council of Forest Ministers, Ottawa, Ontario.

Kapeller, S.S., Lexer, M.J., Geburek, T., Hiebl, J. & Schueler, S. (2012) Intraspecific variation in climate response of Norway spruce in the eastern Alpine range: Selecting appropriate provenances for future climate. *Forest Ecology and Management*, **271**, 46–57.

Köble, R. & Seufert, G. (2001) *Novel maps for forest tree species in Europe. Proceedings of the 8th European Symposium on the Physico-Chemical Behaviour of Air Pollutants: "A Changing Atmosphere!,"* pp. 17–20. Data set available online at: http://afoludata.jrc.ec.europa.eu/index.php/dataset/detail/66, Torino.

Koskela, J., Lefèvre, F., Schueler, S., Kraigher, H., Olrik, D.C. & Hubert, J. (2013) Translating conservation genetics into management : Pan-European minimum requirements for dynamic conservation units of forest tree genetic diversity. *Biological Conservation*, **157**, 39–49.

Ledig, F.T. (1986) Conservation strategies for forest gene resources. *Forest Ecology and Management*, **14**, 77–90.

Leech, S.M., Almuedo, P.L. & Neill, G.O. (2011) Assisted Migration : Adapting forest management to a changing climate. *BC Journal of Ecosystems and Management*, **12**, 18–34.

Lefèvre, F., Koskela, J., Hubert, J., Kraigher, H., Longauer, R., Olrik, D.C., Schueler, S., Bozzano, M., Alizoti, P., Bakys, R., Baldwin, C., Ballian, D., Black-Samuelsson, S., Bednarova, D., Bordács, S., Collin, E., de Cuyper, B., de Vries, S.M.G., Eysteinsson, T., Frýdl, J., Haverkamp, M., Ivankovic, M., Konrad, H., Koziol, C., Maaten, T., Notivol Paino, E., Oztürk, H., Pandeva, I.D., Parnuta, G., Pilipovič, A., Postolache, D., Ryan, C., Steffenrem, A., Varela, M.C., Vessella, F., Volosyanchuk, R.T., Westergren, M., Wolter, F., Yrjänä, L. & Zariŋa, I. (2013) Dynamic conservation of forest genetic resources in 33 European countries. *Conservation biology: the journal of the Society for Conservation Biology*, **27**, 373–84.

Lindner, M., Maroschek, M., Netherer, S., Kremer, A., Barbati, A., Garcia-Gonzalo, J., Seidl, R., Delzon, S., Corona, P., Kolstr??m, M., Lexer, M.J. & Marchetti, M. (2010) Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. *Forest Ecology and Management*, **259**, 698–709.

Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B. & Ackerly, D.D. (2009) The velocity of climate change. *Nature*, **462**, 1052–1055.

Mazziotta, A., Triviño, M., Tikkanen, O.P., Kouki, J., Strandman, H. & Mönkkönen, M. (2015) Applying a framework for landscape planning under climate change for the conservation of biodiversity in the Finnish boreal forest. *Global Change Biology*, **21**, 637–651.

McCullagh, P. & Nelder, J.A. (1989) *Generalized linear models*, (ed. by C. and Hall).

Meier, E.S., Lischke, H., Schmatz, D.R. & Zimmermann, N.E. (2012) Climate, competition and connectivity affect future migration and ranges of European trees. *Global Ecology and Biogeography*, **21**, 164–178.

Morgante, M., Felice, N. & Vendramin, G.G. (1998) *Analysis of hypervariable chloroplast microsatellites in* Pinus halepensis *reveals a dramatic bottleneck. Tools for Screening Biodiversity: Plants and animals* (ed. by A. Karp), P.G. Isaac), and D.S. Ingram), pp. 407–412. Chapman & Hall, London.

Moritz, C. (1994) Defining evolutionarily-significant-units for conservation. *Trends in Ecology & Evolution*, **9**, 373–375.

Mouillot, F., Rambal, S. & Joffre, R. (2002) Simulating climate change impacts on fire frequency and vegetation dynamics in a Mediterranean ecosystem. *Global Change Biology*, **8**, 423–437.

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

Nicotra, A.B., Beever, E.A., Robertson, A.L., Hofmann, G.E. & O'Leary, J. (2015) Assessing the components of adaptive capacity to improve conservation and management efforts under global change. *Conservation Biology*, **29**, 1268–1278.

Normand, S., Ricklefs, R.E., Skov, F., Bladt, J., Tackenberg, O. & Svenning, J.C. (2011) Postglacial migration supplements climate in determining plant species ranges in Europe. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 3644–3653.

Pausas, J.G. (2004) Changes in fire and climate in the eastern Iberian Peninsula (Mediterranean Basin). *Climatic Change*, **63**, 337–350.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Rehfeldt, G.E., Worrall, J.J., Marchetti, S.B. & Crookston, N.L. (2015) Adapting forest management to climate change using bioclimate models with topographic drivers. *Forestry*, 1–12.

Rodríguez-Quilón, I., Santos-del-Blanco, L., Serra-Varela, M.J., Koskela, J., González-Martínez, S.C. & Alía, R. (2016) Capturing neutral and adaptive genetic diversity for conservation: maritime pine as a case study. *Ecological Applications*, **accepted**.

Santos-del-Blanco, L., Bonser, S.P., Valladares, F., Chambel, M.R. & Climent, J. (2013) Plasticity in reproduction and growth among 52 range-wide populations of a Mediterranean conifer: Adaptive responses to environmental stress. *Journal of Evolutionary Biology*, **26**, 1912–1924.

Savolainen, O., Pyhäjärvi, T. & Knürr, T. (2007) Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 595–619.

Schueler, S., Falk, W., Koskela, J., Lefèvre, F., Bozzano, M., Hubert, J., Kraigher, H., Longauer, R. & Olrik, D.C. (2014) Vulnerability of dynamic genetic conservation units of forest trees in Europe to climate change. *Global Change Biology*, **20**, 1498–1511.

Schueler, S., Kapeller, S., Konrad, H., Geburek, T., Mengl, M., Bozzano, M., Koskela, J., Lefèvre, F., Hubert, J., Kraigher, H., Longauer, R. & Olrik, D.C. (2013) Adaptive genetic diversity of trees for forest conservation in a future climate: a case study on Norway spruce in Austria. *Biodiversity and Conservation*, **22**, 1151–1166.

Serra-Varela, M.J., Grivet, D., Vincenot, L., Broennimann, O., Gonzalo-Jiménez, J. & Zimmermann, N.E. (2015) Does phylogeographic structure relate to climatic niche divergence? A test using maritime pine (*Pinus pinaster* Ait.). *Global Ecology and Biogeography*, **24**, 1302–1313.

Sykes, M.T., Prentice, I.C. & Cramer, W. (2009) A bioclimatic model for the potential distributions of north European tree species under present and future climates. *Journal of Biogeography*, **23**, 203–233.

Thuiller, W., Lavorel, S., Sykes, M.T. & Araújo, M.B. (2006) Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. *Diversity and Distributions*, **12**, 49–60.

Voltas, J., Chambel, M.R., Prada, M.A. & Ferrio, J.P. (2008) Climate-related variability in carbon and oxygen stable isotopes among populations of Aleppo pine grown in common-garden tests. *Trees - Structure and Function*, **22**, 759–769.

Wang, T., Hamann, A., Yanchuk, A.D., O'Neill, G.A. & Aitken, S.N. (2006) Use of response functions in selecting lodgepole pine populations for future climates. *Global Change Biology*, **12**, 2404–2416.

Watson, J.E.M., Iwamura, T. & Butt, N. (2013) Mapping vulnerability and conservation adaptation strategies under climate change. *Nature Climate Change*, **3**, 989–994.

Supporting Information

Additional Supporting Information may be found in the annex section.

Appendix S1. Details on *Pinus halepensis* Mill. genetic clades: Q membership values and description of the clades.

> **Table S1.** Estimated membership coefficients (*Q*) for each population. Q values above 0.80 for one clade (K) correspond with pure populations for that specific K. Q coefficients above 0.25 for one K indicate presence of that K into the population. Colors correspond to overall genetic clusters eventually defined for the study.

Appendix S2. Evaluation scores obtained in the Species Distribution Models

> **Table S2.** True Skill Statistic (TSS) scores obtained by the individual statistic algorithms models of the eight genetically defined clades of *Pinus pinaster* Ait. (G1-pin to G8-pin)
>
> **Table S3.** True Skill Statistic (TSS) scores obtained by the individual statistic algorithms models of the seven genetically defined clades *Pinus halepensis* Mill. (G1-hal to G7-hal).
>
> **Table S4.** True Skill Statistic (TSS) scores, Cutoff values, Sensitivity and Specificity values of the ensemble models of the eight genetically defined clades of *Pinus pinaster* Ait (G1-pin to G8-pin)
>
> **Table S5.** True Skill Statistic (TSS) scores, Cutoff values, Sensitivity and Specificity values of the ensemble models of the seven genetically defined clades *of Pinus halepensis* Mill. (G1-hal to G7-hal)

Appendix S3. Conservation guidelines and maps of *Pinus pinaster* Ait. and *Pinus halepensis* Mill. infraspecific clades: Distribution, future suitability maps and conservation proposals.

> **Fig. S1 – Fig. S8:** *Pinus pinaster* Ait.
>
> **Fig. S9 – Fig. S15:** *Pinus halepensis* Mill.
>
> **Table S6**. Areas of to the territories proposed for the different conservation strategies within each clade for both *P. pinaster* Ait. and *P. halepensis* Mill.

Appendix S4. Dynamic Conservation Units: Classification within clades and Exposure assessment

Table S7 *Pinus pinaster* Ait. Dynamic conservation Units

Table S8 *Pinus halepensis* Mill. Dynamic conservation Units

Table S9: *Pinus pinaster* Ait. conservation relevant population groups

Chapter 3

Incorporating exposure to pitch canker disease to support management decisions of *Pinus pinaster* Ait. in the face of climate change

Incorporating exposure to pitch canker disease to support management decisions of *Pinus pinaster* Ait. in the face of climate change

Serra-Varela *et al.* in revision in PLOSONE

Serra-Varela, M.J.[1,2,3] maje.serra@gmail.com

Alía, R. [2,3] alia@inia.es

Pórtoles, J.[4] fic@ficlima.org

Gonzalo, J. [2,3] jgonzalo@pvs.uva.es

Soliño, M. [3,5] solino.mario@inia.es

Grivet, D. [2,3] grivet@inia.es

Raposo, R.[3] raposo@inia.es

1 – University of Valladolid, Department of Plant Production and Forest Resources, Avda. de Madrid 44, 34004 Palencia, Spain

2 - Sustainable Forest Management Research Institute, INIA- University of Valladolid, Avda. Madrid 44, 34004 Palencia, Spain.

3 – INIA, Forest Research Centre, Department of Forest Ecology and Genetics, Carretera A Coruña km 7.5, 28040 Madrid, Spain.

4 – Fundación para la Investigación del Clima, FIC. Gran Vía 22 – Dpdo., 28013, Madrid, Spain

5 - INIA, Forest Research Centre, Department of Forestry and Management, Carretera A Coruña km 7.5, 28040 Madrid, Spain.

Corresponding author: María Jesús Serra-Varela; maje.serra@gmail.com; Crra de la Coruña km 7.5, 28040, Madrid Phone: 91 347 39 04; Fax 91 357 22 93

Abstract

Climate change and its impacts on abiotic and biotic factors are largely affecting forest ecosystems resulting in shifts in their distribution ranges and in increased infections and biological invasions. Thus, it is necessary to assess the magnitude of climatic changes (*i.e.* exposure) on both abiotic and biotic terms, in order to guide management for adaptation and enhance species' capacities to cope with new climatic conditions. In this work, we integrated both factors in order to assess the exposure of the conifer *Pinus pinaster* Ait. to climate change, and to obtain useful information to guide its management. We approached abiotic exposure by means of risk of habitat loss, whereas we evaluated biotic exposure based on the future climatic suitability of the pitch canker, a pine disease caused by *Fusarium circinatum* Nirenberg & O'Donnell. As we aimed to obtain useful and reliable information for managers, we constrained our study area to the Spanish Iberian Peninsula where we disposed of accurate climate and occurrence databases. While *P. pinaster* is widely distributed across the study area, the disease has only been detected in its north and north-western edges. We fit species distribution models for the current distribution of the conifer and the disease. Then, we projected these models into 18 different future climate simulations representative of 2050 and created two future suitability maps (for the pine and for the disease individually) based on the level of agreement among future projections. Finally, we used future suitability maps to assess whether current populations of *P. pinaster* were exposed to climate change in both abiotic and biotic terms. We found that almost the entire distribution of *P. pinaster* in the Spanish Iberian Peninsula was subjected to abiotic exposure likely to be driven by the predicted increase in drought events in the future. On the other hand, likely suitable areas in the future for the disease concentrated in the north of the study area whereas we found a reduction in its suitable habitat in its north-western distribution. Based on these results, with the objective of enhancing maritime pine's capacity to cope with two major stresses, we recommend setting up breeding programmes in highly exposed and productive populations, while silvicultural methods and monitoring should be applied in those less productive but still exposed populations.

Keywords: maritime pine, *Fusarium circinatum*, Species Distribution Models, SDMs, vulnerability

1. Introduction

Anthropogenic climate change largely affects forest ecosystems, claiming for the need of management actions for adaptation. The development of management plans for adaptation requires an assessment of exposure, *i.e.* an evaluation of the magnitude of climate change (see Dawson *et al.,* 2011), integrating its impacts on abiotic and biotic factors (Bolte *et al.*, 2009). The alterations of abiotic factors due to climate change, such as increased intensity and/or duration of droughts (mid-latitudes) and ascending global mean temperatures (IPCC, 2013), have led to increasing tree mortality (Allen *et al.,* 2010) and northwards shifts of the distributions of many species (Walther *et al.*, 2012). Because of these major climate-related shifts, abiotic exposure to climate change has been commonly addressed in literature (*e.g.* Coops & Waring, 2011; Johnston *et al.,* 2009; Schueler *et al.,* 2014). In contrast, there have been few studies addressing how pests and pathogens are responding to climate change and their effect on hosts (but see Chakraborty, 2013; Pautasso *et al.,* 2015; Sturrock *et al.,* 2011), although the importance of climate change on biotic factors is progressively becoming evident. For instance, changes in temperatures and seasonal precipitation are responsible for increased infection of several species of *Phytophthora* in European forests (Jung, 2009; Redondo *et al.*, 2015) and of *Dothistroma pini* and *D. septosporum* in the northern hemisphere (Welsh *et al.*, 2014). Climate and global change are also responsible for increasing biological invasions (Pautasso *et al.*, 2010; Sturrock *et al.*, 2011) which constitute new threats for forests. In fact, only in Europe during the last four decades, forest pathogen invasions have grown exponentially, with introductions mainly from North America and recently, from Asia (Santini *et al.*, 2013). In most situations, these invasions had an important consequence on native tree species: *Ophiosthoma novo-ulmi* in Western Europe had a devastating impact on mature elm trees (*Ulmus minor*) in the 1970s (Anderson *et al.*, 2004), while *Hymenoscyphus fraxineus* has caused extensive ash (*Fraxinus excelsior*) dieback throughout Europe since the 1990s (Bakys *et al.*, 2009). Still, biotic exposure is rarely contemplated when assessing vulnerability to climate change (but see Heikkinen *et al.,* 2007; Preston *et al.,* 2008 for examples with birds), thus limiting the scientific basis of management for adaptation, which would promote resilience of forest ecosystems to climate change (Bolte *et al.*, 2009).

Assessing the impact of climate change on a potential pathogen invasion is challenging since climate change will affect the pathogen, the host and the dynamic of their interaction. There are at least two requirements that need to be fulfilled for a successful pathogen invasion, namely that the pathogen is able to infect the host species and that the pathogen is able to survive in the same environmental conditions as the host species. Evaluating the first prerequisite is difficult as we generally lack empirical information and because climate change may also affect species interactions (Tylianakis *et al.*, 2008). The second prerequisite can be more easily evaluated and provides a first insight into biotic exposure (*e.g.* Giljohann *et al.,* 2011). Approaches that integrate both, abiotic and biotic exposure to climate change constitute a solid tool to effectively guide forest management and breeding programmes directed to increase forest capacity to cope with climate change, and to guarantee that they maintain their essential role of providing services for the society (see FAO, 2006).

In this framework species distribution models (SDMs - see Guisan & Zimmermann, 2000 for a detailed description) provide a useful tool as they evaluate a species' habitat suitability under different climatic scenarios based on known occurrence records and on environmental variables. Abiotic exposure can be approached by the probability that the habitat of a species will remain suitable in the future and, similarly, we can assess whether a territory is likely or not to become suitable for new pests and diseases affecting the target species and consequently assess its potential biotic exposure (see Baxter & Possingham, 2011; Giljohann *et al.,* 2011; Thuiller *et al.,* 2005 for some examples).

Here, we evaluated exposure to climate change of *Pinus pinaster* Ait. (maritime pine), an iconic, ecologically and economically important Mediterranean conifer, integrating both, abiotic and biotic factors. We assessed abiotic exposure by estimating risk of habitat loss in future climate predictions representative of 2050. We incorporated biotic exposure by considering pitch canker disease, an emergent forest disease of *Pinus* species caused by the fungus *Fusarium circinatum* Nirenberg & O'Donnell, as a potential biotic threat for *P. pinaster*. This disease was first detected in Europe in pine nurseries in 2005 (Landeras *et al.*, 2005). Since then, it has been reported in *P. radiata* plantations in Northern Spain (Pérez-Sierra *et al.*, 2007; Iturritxa *et al.*, 2011), and only rarely in *P. pinaster* (see Iturritxa *et al.,* 2013). Nevertheless, the scientific community is concerned that disease incidence in *P. pinaster* could be largely increased due to global change, with important ecologic and economic consequences, in the case of larger infected areas or of higher inoculate concentration.

A combined approach assessing habitat suitability for both *P. pinaster* and the disease has still not been addressed although previous studies have evaluated it separately (see Serra-Varela *et al.,* 2015,in revision for *P. pinaster* and Baker *et al.,* 2010; Ganley *et al.,* 2009; Iturritxa *et al.,* 2015; Möykkynen *et al.,* 2014; Watts *et al.,* 2011 for pitch canker disease). From these, only Serra-Varela *et al.,* (in revision), Möykkynen *et al.* (2014) and Watt *et al.* (2011) assessed suitability in the future. Nevertheless, given their large scale input climatic and distribution databases, these constitute broad scale approaches aiming to obtain general trends. Here, we aim to fulfil the accuracy and resolution requirements needed for local forest management. Consequently, we constrained our study area to the Spanish Iberian Peninsula for which we have specific sources of information (*i.e.* data from the Spanish Meteorological Agency – AEMET - with over 5000 meteorological stations available and from the Spanish National Forest Inventory in which sampling is performed at a 1Km grid). A solid and high resolution assessment of abiotic and biotic exposure of *P. pinaster* is of particular interest for managers, who can opt for distinct practices in order to enhance this economically important pine to cope with these stresses: breeding programmes can be applied for highly exposed and productive populations, while particular silvicultural methods (Prieto-Recio *et al.*, 2015) and monitoring could be set up for less productive but still exposed populations.

2. Materials and Methods

2.1. Occurrence data

SDMs require an occurrence dataset including presence records (for all statistical algorithms) and, in general, also absence, pseudo-absence or background records of the target species (see Pearson, 2010 for more details) to base their predictions upon. Nevertheless, the difficulty associated to detect real absence records translates into a generalized use of pseudo-absence or background records. Thus, we obtained an occurrence dataset composed of presence and pseudo-absence records for the pine species (*P. pinaster*) and the pitch canker disease along the study area (*i.e.* Spanish Iberian Peninsula). In this section, we explain the method for selecting presences but given that the number of pseudo-absences and their selection technique strongly depends on the specific algorithm selected (see Barbet-Massin *et al.*, 2012 for details) it will be explained in the SDM section along with other algorithm specifications.

In the case of *P. pinaster*, we employed the third Spanish National Forestry Inventory (NFI) developed between 1997-2007 based on a 1 km grid. We selected as presences those plots where natural and seed-born populations of *P. pinaster* were reported as one among the three major species, which after removing duplicates, led to a set of 6081 plots. To avoid possible misguiding results due to the inclusion of non-native populations in the analysis, we further eliminated those plots not included within the native distribution of *P. pinaster* (assessed from Alía *et al.*, 2009). This reduced the number of selected plots to 2971, representative of the native distribution of *P. pinaster* along the study area and we used the central coordinates of the plots as presences records. We considered the rest of the study area, not fulfilling these criteria, as potential pseudo-absence records.

Concerning pitch canker disease, we used data from a survey performed by different Regional Authorities in Spain and collected by the Spanish Ministry of Agriculture, Food and Environment concerning the period 2006-2012. We obtained 159 municipalities, located in the north-western side of the Iberian Peninsula where at some point during that period a disease outbreak was declared. All pitch canker disease outbreaks were reported in *Pinus radiata* D.Don. In order to select adequate coordinates to represent presence records, we divided the study area in a 1 Km grid and selected the 1444 grid cells that were simultaneously (i) within the positive-detected municipality borders and (ii) within the host species distribution (*i.e. P. radiata* obtained from the NFI). Finally, to prevent the municipality size from affecting model outputs (as bigger municipalities tended to have higher number of disease presences included in the models) we considered a maximum of 10 presences per municipality *i.e.* the mean number of potential disease presences per municipality. Thus, in those municipalities enclosing less than 10 potential presences all records were included within the definitive presence data set, whereas in those municipalities where potential presences were larger, we randomly selected 10 records to be included within the final presence dataset. The definitive presence dataset reckoned 943 records. Similarly to the case of *P. pinaster*, we considered all the rest of the study area, as potential pseudo-absence records.

2.2. Bioclimatic data

Global data bases do not fulfill the requirements in accuracy or resolution to support local forest management. Indeed, WORDLCLIM (Hijmans *et al.*, 2005) has already been reported as problematic in the Spanish Iberian Penisula (see Bedia *et al.,* 2013) as its interpolations are based on 142 clustered distributed, meteorological stations, among which highest altitudes are barely represented (see Fig. S1 in Appendix S1). To avoid this problem, here, we used 5053 meteorological stations with observed daily precipitation data and 1830 with observed daily maximum and minimum temperature data from AEMET, covering the period between 1950-2000 to obtain accurate and high resolution bioclimatic surfaces (see Fig. S1 in Appendix S1 for a comparison between WORDLCLIM and AEMET meteorological stations). First, we calculated monthly variables as monthly accumulated precipitation and as monthly means for maximum and minimum temperature. Then, we interpolated these monthly variables by means of Thin Plate Splines (TPS) (Hutchinson, 1991) using elevation as independent co-variable to obtain continuous surfaces (1 Km grid cell) across the study area. We selected TPS as our interpolation method as it has performed well in previous comparative tests of multiple interpolation techniques (Hartkamp *et al.,* 1999; Jarvis & Stuart, 2001), it has been widely used previous studies - including WORLDCLIM (Hijmans *et al.*, 2005) and others such as *New et al.* (1999) – and because it is computationally efficient and easy to run. Finally, we calculated the 19 bioclimatic variables proposed by WORLDCLIM (Hijmans *et al.*, 2005) (BIO1-BIO19; see Table 1).

As additional variables, we also included distance to the coast (dist_coast) and two topographic variables namely elevation and slope, both derived from GTOPO30 model (courtesy of the U.S. Geological Survey) as these have been reported as relevant in determining the distribution of the disease (Rosso & Hansen, 2003; Iturritxa *et al.*, 2015). Table 1 provides for a summary of the variables tested.

In order to avoid multicollinearity effects, we retained variables with Pearson correlations lower than 0.60 (see Table 1 for more information) as the use of simple methods based on rules of thumb has proved to be as effective as more complicated methods (Dormann *et al.*, 2013). Among highly correlated variables we kept the one with highest explained deviance scores (D^2) when individually fitted in a Generalized Linear Model (GLM - McCullagh and Nelder, 1989). We avoided the use of BIO8 and BIO9 as the steep gradient shown by these variables, in which very often two adjacent cells are characterized by extremely different values within the study area for no obvious reason, may lead to artefacts in the SDM output maps. Finally, we checked for possible collinearity by performing a Variance Inflation Factor (VIF; Belsley, 1991; Hair Jr *et al.*, 1995), ensuring that all VIF values were below 5. The final sets of relevant weakly correlated variables to build SDMs were (i) BIO4 - Temperature Seasonality, BIO12 - Annual Precipitation, BIO17 - Precipitation of Driest Quarter and elevation for *P. pinaster* and, (ii) BIO4 - Temperature Seasonality, BIO6 - Mean Temperature of Coldest Month, BIO12 - Annual Precipitation, BIO17 - Precipitation of Driest Quarter and Distance to the coast for the disease. For both, the pine and the disease, we also calculated mean and standard deviations values of the environmental variables separately for presences and absences to get insights about their currently inhabited habitat conditions.

Table 1: Complete list of environmental variables tested as candidates to be included in species distribution models (SDMs) for the pitch canker disease and *Pinus pinaster* Ait. D^2 indicates the explained deviance score obtained when individually fitting the variable in a Generalized Linear Model (GLM). Similarly coloured rows group highly correlated variables (Pearson correlation > 0.60) while non-coloured ones indicate non-correlated variables. Variables in bold show finally selected variables.

TYPE OF VARIABLE	VARIABLE	EXPLANATION	DATA SOURCE	D^2 P. pinaster	D^2Pitch canker
CLIMATIC	BIO1	Annual Mean Temperature	AEMET	0.12	0.24
	BIO2	Mean Diurnal Range: Mean of monthly (max temp - min temp)	AEMET	0.01	0.43
	BIO3	Isothermality (BIO2/BIO7)	AEMET	0.01	0.32
	BIO4	**Temperature Seasonality (standard deviation)**	AEMET	**0.04**	**0.57**
	BIO5	Max Temperature of Warmest Month	AEMET	0.14	0.50
	BIO6	**Min Temperature of Coldest Month**	AEMET	0.08	**0.24**
	BIO7	Temperature Annual Range (BIO5-BIO6)	AEMET	0.04	0.57
	BIO8	Mean Temperature of Wettest Quarter	AEMET	0.04	0.11
	BIO9	Mean Temperature of Driest Quarter	AEMET	0.08	0.39
	BIO10	Mean Temperature of Warmest Quarter	AEMET	0.15	0.36
	BIO11	Mean Temperature of Coldest Quarter	AEMET	0.10	0.22
	BIO12	**Annual Precipitation**	AEMET	**0.06**	**0.49**
	BIO13	Precipitation of Wettest Month	AEMET	0.01	0.42
	BIO14	Precipitation of Driest Month	AEMET	0.13	0.46
	BIO15	Precipitation Seasonality (Coefficient of Variation)	AEMET	0.01	0.07
	BIO16	Precipitation of Wettest Quarter	AEMET	0.01	0.43
	BIO17	**Precipitation of Driest Quarter**	AEMET	**0.21**	**0.49**
	BIO18	Precipitation of Warmest Quarter	AEMET	0.19	0.47
	BIO19	Precipitation of Coldest Quarter	AEMET	0.02	0.40
TOPOGRAPHIC	**Elevation**	**Elevation above the sea level (m)**	G30TOPO	**0.17**	0.20
	Slope		G30TOPO	0.01	0.09
	Dist_coast	**Distance to the coast**	G30TOPO	0.04	**0.40**

Future climatic variables (monthly accumulated precipitation and monthly maximum and minimum temperature) representative of 2050 (average for 2041 – 2060) were obtained from nine of the most recent Global Climate Models (GCMs) that have been used in the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment report (IPCC, 2013) (see Table 2). As future projection simulations, we used the two different Representative Concentration Pathways (RCP), namely RCP4.5 (medium emission scenario) and RCP8.5 (high emission scenario) (Taylor *et al.*, 2009) totalling to 18 different future climate predictions.

GCMs, which have a coarse resolution (*ca.* 2 degrees), were transformed to a local scale using a two-step analogue statistical downscaling method developed by Ribalaygua *et al.,* (2013). The first step is an analogue approach (Zorita & Von Storch, 1999) in which the *n* most similar days to the day to be downscaled are selected by using four different meteorological large-scale fields. The second step differed depending on the target variable: (i) precipitation was calculated by re-assigning the calculated amounts using an empirical distribution function; and (ii) temperature was obtained by using a multiple linear regression analysis using the *n* most analogous days selected. We followed this procedure for each AEMET meteorological station and for each target day, and then averaged the resulting daily simulations to obtain monthly means. Further details of the methodology are described in Ribalaygua *et al.,* (2013).

We obtained a systematic error when comparing the simulated data from climate models with the observed data from reference time series due to the inherent downscaling methodology error and to the inner GCM error (which usually incorporates a bias over the data). Thus, to improve our simulations, future climate projections were corrected according to a parametric quantile–quantile method (Monjo *et al.*, 2014). Then, we interpolated the obtained variables *i.e.* future monthly accumulated precipitation and future montly maximum and minimum temperatures, across the entire study area following the same procedure as employed in current climate interpolations *i.e.* TPS. Finally, we calculated WORLDCLIM bioclimatic variables representative of 2050 (BIO1 to BIO19; see Table 1).

Table 2: Global Climate Models (GCMs) employed to obtain future climate predictions representative of 2050

MODEL	INSTITUTION	COUNTRY	RESOLUTION (LongXLat)
BCC-CSM1-1	Beijing Climate Center (BCC), China Meteorological Administration	China	2·8 × 2·8°
CanESM2	Canadian Centre for Climate Modelling and Analysis (CC-CMA)	Canada	2·8 × 2·8°
CNRM-CM5	Centre National de Recherches Meteorologiques/Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique (CNRM-CERFACS)	France	1·4 × 1·4°
GFDL-ESM2 M	Geophysical Fluid Dynamics Laboratory (GFDL)	United States	2 × 2·5°
HADGEM2-CC	Met Office Hadley Centre (MOHC)	UK	1·87 × 1·25°
MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Atmosphere and Ocean Research Institute (AORI), and National Institute for Environmental Studies (NIES)	Japan	2·8 × 2·8°
MPI-ESM-MR	Max Planck Institute for Meteorology (MPI-M)	Germany	1·8 × 1·8°
MRI-CGCM3	Meteorological Research Institute (MRI)	Japan	1·2 × 1·2°
NorESM1-M	Norwegian Climate Centre (NCC)	Norway	2·5 × 1·9°

2.3. Species Distribution Models

We used generalized additive models (GAM; Hastie & Tibshirani, 1990) by which we fit response curves using a non-parametric smoothing function to model the relationship between the environmental variables and each occurrence dataset. Models were processed in BIOMOD (Thuiller *et al.*, 2009) using the package "biomod2" (default settings for the model) in the R statistical software environment (R Development Core Team, 2016). As recommended for GAM by Barbet-Massin *et al.* (2012), we randomly selected 10000 pseudo-absences and same weight was given to presences and absences. Finally, we performed a five independent 70-30 training-evaluating subsets of the data for model evaluation and assessed model performance as the mean True Skill Statistic (TSS; Allouche *et al.,* 2006). Thus, we obtained a probabilistic prediction of habitat suitability for each climatic data set (current climate and 18 future climate predictions) and for each, the pine and the disease. These were converted to binary presence-absence maps by defining thresholds that maximized TSS values.

We analysed variable importance following the method available in the package "biomod2": First, the model was trained with the selected environmental variables and predictions were performed (reference predictions). Then, one environmental variable was changed to random values and new predictions were done. This process was repeated four times with each environmental variable included in the model. Finally, we calculated Pearson´s correlations - which range between 0 (no correlation) and 1 (maximum correlation) - between reference predictions and the new ones. The final score for each environmental variable is provided as the difference between 1 and mean correlation among the four repetitions performed for each variable so that higher values indicate higher importance.

2.4 Future suitability maps

For the pine and the disease separately, we obtained 18 future binary projections which classify each grid cell of the study area as suitable or unsuitable in 2050. In order to incorporate the uncertainty derived from the wide range of future climate predictions available avoiding the use of an average prediction, we followed the methodology proposed in Serra-Varela *et al.* (in revision). Accordingly to their work, we used the number of future climate predictions projecting suitability in one specific cell as an indicator of the degree of agreement among models that the future habitat will be suitable in that cell . Thus, we combined the 18 binary projections and obtained a map in which values could possibly range from 0 (none of the future climate predictions was projected to be suitable) to 1.0 (all future climate predictions tested were projected to be suitable) with higher scores indicating higher agreement of suitable habitat in the future. Finally we defined three different future suitability categories: "likely suitable" with suitability scores >0.7 (suitable habitat for more than 70% of future projections), "uncertain" with a suitability score of 0.36-0.7, and "likely unsuitable" with suitability scores <0.36. We performed this analysis for each species individually.

Our future suitability maps rely on the assumption of niche conservatism during a 30 year timespan (*i.e.* a constant relationship between each species and climate). This is expectable in the case of *P. pinaster*, as it is long-lived tree with long rotation times (from 30 to almost 100 years) and thus, only one generation should be involved during this period. Furthermore, *F. circinatum,* the fungus responsible for the pitch canker disease, reproduces asexually, and evidence supports that sexual reproduction does not occur (Berbegal *et al.*, 2013). In Spain, there is a clonal population structure with two genotypes widely distributed (Berbegal *et al.*, 2013) but it is unknown how climate change will influence the growth and reproduction of any of these genotypes and their interaction with the host.

2.5 Abiotic and biotic exposure for *P. pinaster* Ait.

We assessed the abiotic and biotic exposure of *P. pinaster* by means of future suitability maps (both, *P. pinaster* and pitch canker disease respectively) and we used the three different suitability categories defined in them, *i.e.* likely suitable, uncertain and likely unsuitable, as an evidence to recommend the most appropriate management to facilitate that populations are able to cope with climate change. For this analysis, we considered the entire distribution of the species *P. pinaster i.e.* native and planted populations and we presented our results following the Spanish regions of identification and utilization (RIUs) of forest reproductive material (García *et al.*, 2001). RIUs delimitate ecologically homogeneous distribution areas and thus, they are expected to group genetically similar populations likely to be locally adapted as well as to differ in their productivity, and thereby in their economic impact on regional economies. Consequently, RIUs are appropriate as management units and provide a perfect framework to guide forest management.

To assess abiotic exposure, we employed *P. pinaster* future suitability map. Likely suitable areas highlighted locations barely exposed where no special management for adaptation to climate change was needed. Contrarily, likely unsuitable areas indicated high abiotic exposure sites where additional measures should be considered such as breeding programmes or silvicultural actions enhancing the species' capacities to cope with climate change. Furthermore, we used the future suitability map of the disease to assess biotic exposure. Likely suitable locations for pitch canker were considered as high biotic exposure areas where urgent actions were needed to improve the capacities of *P. pinaster* to avoid possible infections, while likely unsuitable locations were considered as low biotic exposure areas where no further actions were needed. Finally, uncertain areas in both future suitability maps, indicated locations where there was not an agreement among future climate predictions and thus, where monitoring is needed in order to see the development of populations and to address its management accordingly.

3. Results

3.1 Bioclimatic data and Species Distribution Models

We found little differences between the mean environmental conditions in presences and pseudo-absences records of *P. pinaster* (see Table 3 and Fig. S2a in Appendix S1), an outcome that was expected given the low explained deviance scores obtained by most of its environmental predictors (see Table 1). Nevertheless, this analysis only considered environmental variables individually and the SDM fitted for *P. pinaster,* which considers altogether the set of selected environmental variables, had an acceptable performance as revealed by its TSS score (0.69) and its sensitivity and specificity values (93.50 and 75.30 respectively).

Contrarily, in the case of pitch canker disease we found large differentiation between the mean environment of presences and pseudo-absences (see Table 3 and Fig S2b in Appendix S1). Its distribution seemed to be constrained by low temperatures (represented by BIO6) as well as low precipitation regimes (BIO12 and BIO17; see Table 3 and Fig. S2b in Appendix S1). Short distance to the coast was also found to be very relevant for the disease's habitat suitability presumably due to the higher relative humidity in these locations, a key factor during the infection stage (Sakamoto & Gordon, 2006; Sakamoto *et al.*, 2007). The SDM of the disease performed very well as assessed by its evaluation scores (TSS = 0.93, sensitivity = 99.26 and specificity = 93.63) revealing a very important role of climate in determining the distribution of the disease. The detected suitable area was located along the north-western side of the Iberian Peninsula (see Fig. 1d & Fig. 1e) which is consistent with the declared infection outbreaks in Spain. We also detected marginal suitability areas along the eastern coast of the Iberian Peninsula, where there are no declared infection outbreaks.

Pseudo-absences are ultimately representing the entire study area so they are very similar for both, the pine and the disease (see Table 3 and Fig. S2 in Appendix S1). All variables included in the SDMs were highly significant for both, the pine and the disease (p-value < 0.0001). As for variable importance (see Table 3), BIO17 obtained by far the highest score in the SDM of *P. pinaster* as compared to the other variables included in the model. Variables in the SDM of the disease obtained evener values scores than in the case of *P. pinaster*, although BIO4 highlighted as the most influent variable. Finally, the threshold values maximizing TSS and thus employed for transforming probabilistic into binary projections were 60.5 and 45.0 respectively for *P. pinaster* and pitch canker disease.

Table 3: Mean values and variable importance, calculated with the package "biomod2" in R Statistical software, of the environmental variables used to fit species distribution models for *Pinus. pinaster* and pitch canker disease. Standard deviation is shown in brackets

Variable	**Mean**				**Variable Importance**	
	P. pinaster		***Pitch canker***		***P. pinaster***	***Pitch canker***
	Presences	Absences	Presences	Absences		
BIO4	638.7 (51.3)	617.8 (84.1)	428.2 (35.9)	618.3 (84.6)	0.19	0.56
BIO6 (ºC)	-	-	3.4 (1.1)	1.1 (2.5)	-	0.36
BIO12 (mm)	694.9 (223.2)	694.2 (345.4)	1452.9 (225.0)	693.2 (341.9)	0.11	0.46
BIO17 (mm)	80.0 (17.3)	81.6 (51.23)	199.2 (46.1)	81.5 (51.2)	0.50	0.45
Dist. Coast (Km)	-	-	17366.3 (12613.8)	129052.6 (86349.8)	-	0.40
Elevation (m)	968.8 (240.6)	685.84 (396.5)	-	-	0.25	-

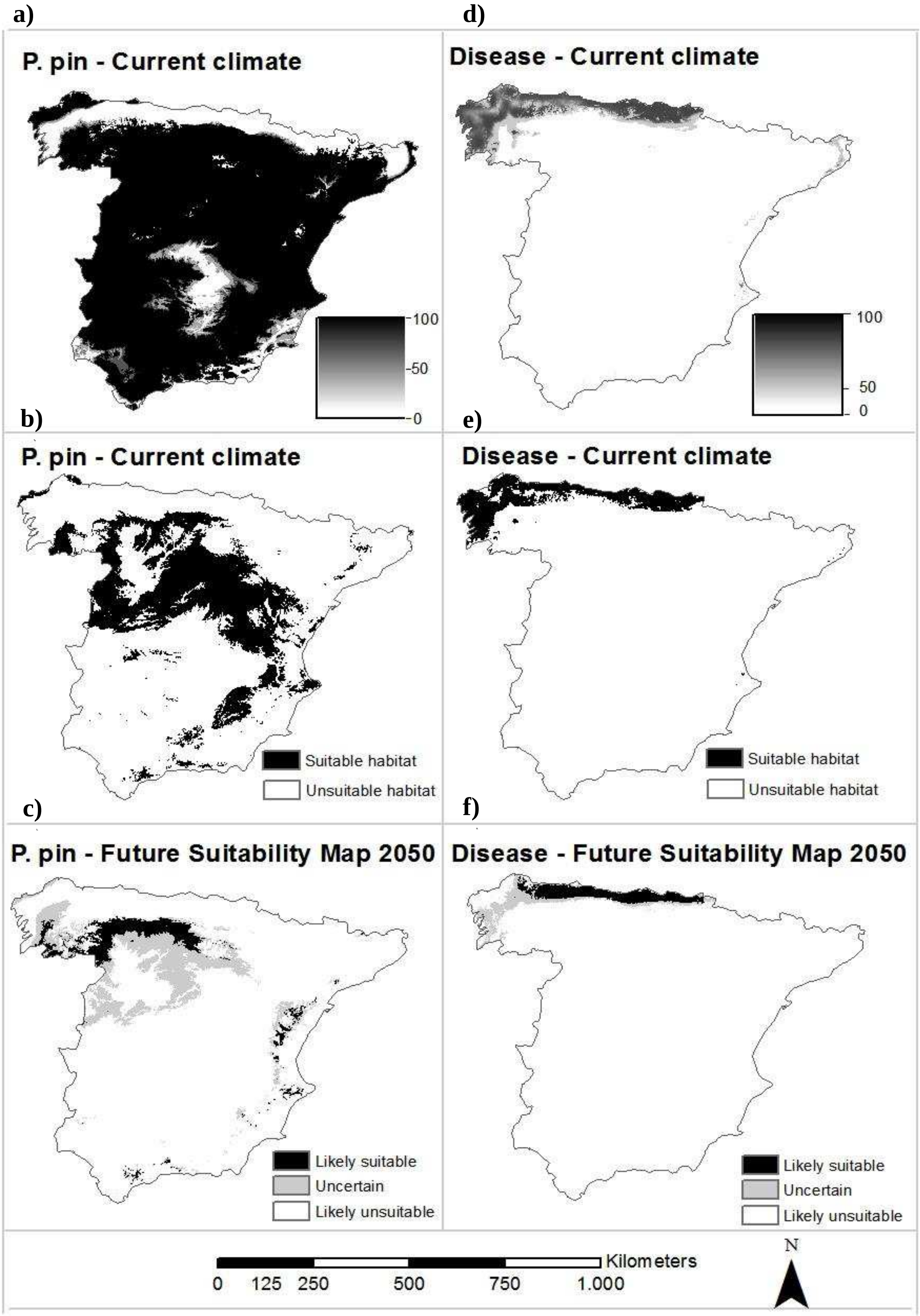


Figure 1: Geographic projections of species distribution models of *Pinus pinaster* (P. pin; a-c) and Pitch canker disease (Disease; d-f). Current climate projections are shown in probabilistic projections - a) and d) - and in binary projections – b) and e). Future suitability maps summarizing 18 future climate predictions are shown in c) and f).

3.2 Future suitability maps

Future suitability maps revealed a reduction in suitable territory for both the pine and the disease (see Fig. 1c & Fig. 1f). In the case of *P. pinaster*, the reduction in habitat suitability was more severe along the central and south-western Iberian Peninsula as most of the likely suitable territory concentrated in the central plateau (RIUs 5, 16 and 17; comparison between Fig. 1b and Fig. 1c). Concerning pitch canker disease, the reduction in habitat suitability mainly occurred in the north-western edge of the study area (RIUs 1 and 2) while the rest of likely suitable areas, was very similar to that detected as currently suitable for the species (comparison between Fig. 1e & Fig. 1f).

3.3 Abiotic and biotic exposure for P. pinaster Ait.

We assessed exposure across all the populations of *P. pinaster* within the study area depending on its own future suitability map and on the future suitability map of its disease (see Fig. 2).

Our results showed that all populations composing the distribution of *P. pinaster* are exposed to a certain degree (but for ca. 8% of the distribution, concentrated around the central plateau and north-western edge of the study area, *i.e.* RIUs number 1, 2, 5, 16 and 17, which were classified as low exposure for both, abiotic and biotic factors). Nevertheless, only *ca* 5 % of the distribution of *P. pinaster* was classified as highly exposed to both abiotic and biotic factors. These populations were located in the northernmost distribution of the species, majorly in the populations from RIUs 3 and 6 (see Fig. 2).

Abiotic exposure affected *P. pinaster* across its entire distribution; in fact, over 50% of its distribution was classified as highly exposed to abiotic factors. Contrarily, areas with high biotic exposure concentrated in the northern edge of the study area which translated into almost 80 % of the distribution classified to be under low biotic exposure.

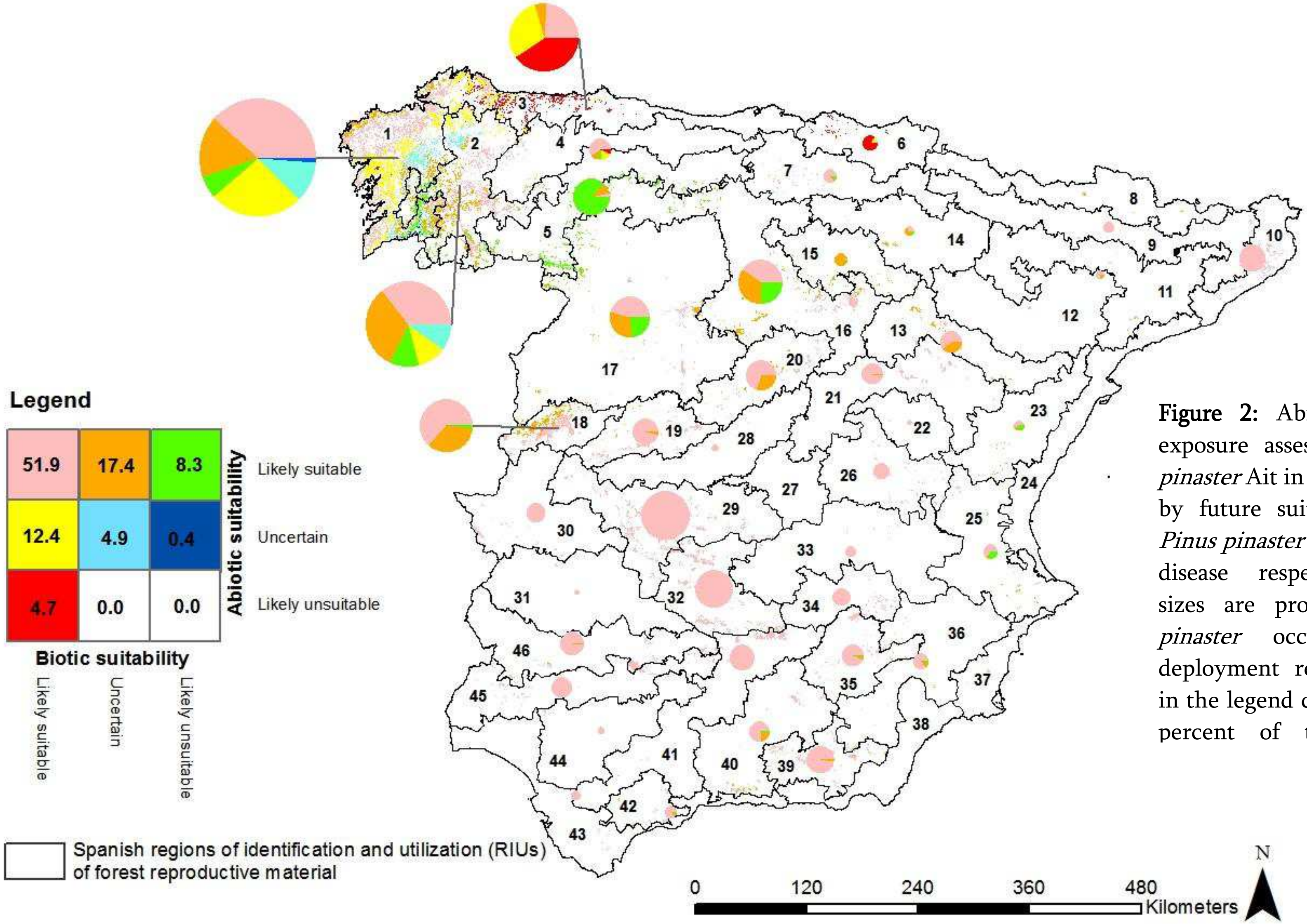


Figure 2: Abiotic and biotic exposure assessment of *Pinus pinaster* Ait in 2050 approached by future suitability maps of *Pinus pinaster* and pitch canker disease respectively. Charts sizes are proportional to *P. pinaster* occupancy within deployment regions. Numbers in the legend correspond to the percent of the distribution

4. Discussion

Designing strategic management plans for adaptation, directed to facilitate that species withstand anthropogenic climate change, requires integrated approaches assessing both, abiotic and biotic exposures (Bolte *et al.*, 2009). These plans are essential for forest owners and managers, as they guide the application of appropriate mitigation measures aiming to guarantee that forests maintain their role in providing their ecological and economical services. Management for adaptation englobes two different strategies, namely (i) breeding programmes incorporating new and adapted genes and, (ii) silvicultural methods directed to enhance local adaptation. Selecting one of them should directly depend on exposure and on the importance of the target population on regional economies. For instance, breeding programmes, which require high investments, are normally restricted to highly productive (and exposed) populations. Here, we integrated abiotic and biotic exposures to climate change of *P. pinaster* in order to support its local forest management.

Overall, the entire distribution of the pine in the Spanish Iberian Peninsula is affected by the impacts of climate change on abiotic factors (but for some exceptions – see RIUs 1, 2, 5, 16 and 17). Abiotic exposure seems to be particularly important along the central and southernmost edge of the Iberian peninsula as all populations in these locations are highly exposed, which is consistent with the results obtained in Serra-Varela *et al.* (in revision). These locations correspond to the rear edge of the distribution of the species and are prone to suffer more strongly the consequences of climate change than others located at northern locations in Europe (Hampe & Petit, 2005). Abiotic exposure is likely to be a consequence of the predicted increased intensity and/or duration of droughts (Mediterranean mid-latitudes IPCC, 2013), as suggested by the importance of BIO17 (Precipitation of Driest Quarter) in determining the species' distribution (see Table 1 & 2). In fact, variables representing drought were also found as relevant in determining the distribution of the pine in previous studies including its entire range *(*Serra-Varela *et al.,* 2015, in revision) and in other local studies (Madrigal-González & Zavala, 2014; Prieto-Recio *et al.*, 2015). Thus, forest management of *P. pinaster*, would benefit from strategies directed to improve the species' capacities to deal with drought stresses. In this direction, breeding programmes are likely to be effective as differences in responses to drought have been attributed to genetic variation (see Gaspar *et al.,* 2013). Nevertheless, because breeding programmes are an expensive procedure they cannot be applied to all populations, and should be restricted to those populations with highest impacts on regional economies, *i.e.* those most productive, which concentrate at the northernmost edge of the distribution (*i.e.* RIUs 1 to 6). For the remaining populations, silvicultural actions directed to fasten adaptation and monitoring should be applied. Among these actions, measures directed at reducing competition for (water) resources, such as decreasing tree density (Prieto-Recio *et al.*, 2015) and favouring unevenness between sizes and ages of the trees in the stand (Madrigal-González & Zavala, 2014), are particularly adequate for *P. pinaster*. In addition, thinning (Prieto-Recio *et al.*, 2015) and extending rotation times - as juveniles are more sensitive to droughts than adults (Madrigal-González & Zavala, 2014) - are other recommended management guidelines for threatened populations.

Biotic exposure to pitch canker disease was assessed by models that captured the relationship between the distribution of the disease and environmental variables. Current suitable locations for the disease concentrated in the north and north-western edge of the study area, which was consistent with previous studies predicting potential global distribution of pitch canker disease on *Pinus* spp (Ganley *et al.*, 2009). Furthermore, we found that future climatic changes are likely to constrain its distribution to the north of the Iberian Peninsula (majorly RIUs 3 and 6), thus reducing biotic exposure on the north western pine populations (RIUs 1 and 2). Breeding programmes enhancing the species' resistance to the pathogen are likely to be successful as differences in responses to *F. circinatum* have been reported to be genetically driven (Vivas *et al.,* 2012). In fact, in a provenance/progeny trial (Elvira-Recuenco *et al.*, 2014), large differences in resistance were found among and within populations suggesting a good host response to the disease through natural or artificial selection, and predicting a successful adaptation of *P. pinaster* to the disease. In addition, high biotic exposure areas coincide with most productive RIUs and consequently investing in breeding programmes would be justified although they must be applied exclusively in exposed populations. Uncertain areas (particularly those from RIUs 1 and 2) should at least be shortly monitored given the important consequences of a disease outbreak and that pine that changes on abiotic factors may interact with biotic factors, *e.g.* stressing the pine and facilitating the infection.

Our *P. pinaster* suitability maps do not pretend to inform about the distribution (present or future) of the species, for which other variables should be included in the model including local adaptation and phenotypic plasticity (Morin & Thuiller, 2009; Valladares *et al.*, 2014), as well as migration (Normand *et al.*, 2011; Hamann & Aitken, 2013). In fact areas of suitable habitat extended largely beyond the current distribution of *P. pinaster* (Figures 1a & 1b) indicating that other factors not related to climate and not considered in our models, such as soils, anthropogenic interactions or intraspecific competition, could be playing an important role in determining the distribution of the pine. Instead, our models aim to assess the relationship between climate and the pine's distribution in order to evaluate whether the pine's current populations are likely to be subjected to abiotic stress in the future.

The obtained suitability maps of pitch canker disease can be used to assess biotic exposure of any other (pine) species affected by this fungus although its incidence or severity would vary according to the species susceptibility. In fact, for this study, we trained our models based on infections detected in *P. radiata* although there is some evidence that suggests that *P. pinaster* is more resistant to the infection, namely (i) the area under disease progress curve (AUDPC) for *P. pinaster* was lower than for *P. radiata* in field experiments (Martinez-Alvarez *et al.*, 2014) and, (ii) the mean wound size for both host species was 27 mm for *P. radiata* and 5 mm for *P. pinaster* in inoculated seedlings (Iturritxa *et al.*, 2012). However, procuring a consistent model of pitch canker disease is challenging given the difficulty associated to obtain a reliable occurrence dataset. In addition, other factors not related to climate such as the distribution of the host species, the different resistance of the pines to the disease and effective prevention plans also influence the current distribution of the disease.

Our work, illustrated with *P. pinaster*, provides an approach to assess abiotic and biotic exposure with regards to climate change. We supply a useful tool for forest managers as both, the resolution and the accuracy of the information employed are adequate for this objective. Nevertheless, other factors could be considered to better characterize the exposure of the species. For instance, fire should be incorporated as an additional abiotic factor as it plays an important role in its biology (Barbéro *et al.*, 1998; Fernandes & Rigolot, 2007), while other destructive pine pests such as the pine processionary moth (Arnaldo & Torres, 2005) or the pine weevil (Zas *et al.*, 2005) should also be included to more fully assess biotic exposure.

Finally, we contribute to forest management in three straight ways: (i) pitch canker future suitability map can be employed to assess biotic exposure of any other susceptible species in the Spanish Iberian Peninsula, (ii) we provide a high resolution and scientifically solid assessment of exposure for *P. pinaster* in the Spanish Iberian Peninsula that can be directly employed in delineating its breeding and management programmes, and (iii) our approach can be easily transferred to any other species.

Acknowledgements

We are grateful to Jesús Martínez Fernández for his collaboration during the analysis stage and to Felipe Bravo-Oviedo for insightful discussion. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for the 5th Coupled Model Intercomparison Project CMIP, and we thank the climate modelling groups (listed in Table 2 of this article) for producing and making available their model output. Thanks are extended to the Spanish Meteorological Agency (AEMET) for providing the climatic observational data, and to Gerardo Sánchez Peña at Spanish Ministry of Agriculture, Food and Environment for providing records of *Fusarium circinatum*.

Funding

We thank the "Scientific research for the adaptation of *Pinus pinaster* Ait. and *Pinus halapensis* Mill. to climate change" project, funded by the Spanish Ministry of Agriculture, Food and Environment, for enabling the collaboration between the different groups involved, as well as the Spanish National Research Plan AdapCon (CGL2011-30182-C02-01), and the TipTree (BiodivERsA-ERANET) which have supported the development of this paper. This study is framed within the project "Integrated Production in Spanish Pine Forests: From the Diversity of Species to Genetic Improvement Programs", funded by the Spanish Ministry of Economy and Competitiveness (RTA2013-00048-C03-01). MJSV was supported by a FPU from the Spanish Ministry of Education, Culture and Sport. D.G. was supported by a Ramón y Cajal fellowship from the Spanish Ministry of Science and Innovation

References

Alía, R., García del Barrio, J.M., S. Iglesias, J.A., Mancha, J., De Miguel, J., Nicolas, L., Pérez, F. & Sánchez de Ron, D. (2009) *Regiones de procedencia de especies forestales en España*, Organismo Autónomo Parques Nacionales, Madrid, Spain.

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

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.

Anderson, P.K., Cunningham, A. a., Patel, N.G., Morales, F.J., Epstein, P.R. & Daszak, P. (2004) Emerging infectious diseases of plants: Pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology and Evolution*, **19**, 535–544.

Arnaldo, P.S. & Torres, L.M. (2005) Spatial distribution and sampling of *Thaumetopoea pityocampa* (Den. & Schiff.) (Lep. Thaumetopoeidea) populations on *Pinus pinaster* Ait. in Montesinho, N. Portugal. *Forest Ecology and Management*, **210**, 1–7.

Baker, R., Candresse, T., Simon, Erzsébet Dormannsné, E., Gilioli, G., Grégoire, J., Jeger, M.J., Karadjova, O.E., Lövei, G., Puglia Porta, A., Rafoss, T., Rossi, V., Schans, J., Schrader, G., Urek, G., Lenteren, J.C. Van & Vloutoglou, I. (2010) Risk assessment of *Gibberella circinata* for the EU territory and identification and evaluation of risk management options. *European Food Safety Authority*, **8**, 1–93.

Bakys, R., Vasaitis, R., Barklund, P., Ihrmark, K. & Stenlid, J. (2009) Investigations concerning the role of *Chalara fraxinea* in declining *Fraxinus excelsior*. *Plant Pathology*, **58**, 284–292.

Barbéro, M., Loisel, R., Quézel, P., Richardson, D.M. & Romane, F. (1998) *Pines of the Mediterranean Basin*. *Ecology and Biogeography of* Pinus (ed. by D.M. Richardson), pp. 153–170. Cambridge University Press, United Kingdom.

Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **3**, 327–338.

Baxter, P.W.J. & Possingham, H.P. (2011) Optimizing search strategies for invasive pests: Learn before you leap. *Journal of Applied Ecology*, **48**, 86–95.

Bedia, J., Herrera, S., Gutiérrez, J.M. & Manuel, J. (2013) Dangers of using global bioclimatic datasets for ecological niche modeling. Limitations for future climate projections. *Global and Planetary Change*, 1–12.

Belsley, D.A. (1991) *Conditioning diagnostics: collinearity and weak data in regression*, Wiley & Sons.

Berbegal, M., Pérez-Sierra, A., Armengol, J. & Grünwald, N.J. (2013) Evidence for multiple introductions and clonality in Spanish populations of *Fusarium circinatum*. *Phytopathology*, **103**, 851–61.

Bolte, A., Ammer, C., Löf, M., Madsen, P., Nabuurs, G.-J., Schall, P., Spathelf, P. & Rock, J. (2009) Adaptive forest management in central Europe: Climate change impacts, strategies and integrative concept. *Scandinavian Journal of Forest Research*, **24**, 473–482.

Chakraborty, S. (2013) Migrate or evolve: options for plant pathogens under climate change. 19(7), 1985-2000. *Global Change Biology*, **19**, 1985–2000.

Coops, N.C. & Waring, R.H. (2011) Estimating the vulnerability of fifteen tree species under changing climate in Northwest North America. *Ecological Modelling*, **222**, 2119–2129.

Dawson, T.P., Jackson, S.T., House, J.I., Prentice, I.C. & Mace, G.M. (2011) Beyond predictions: biodiversity conservation in a changing climate. *Science*, **332**, 53–58.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., Mcclean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D. & Lautenbach, S. (2013) Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 027–046.

Elvira-Recuenco, M., Iturritxa, E., Majada, J., Alía, R. & Raposo, R. (2014) Adaptive potential of maritime pine (Pinus pinaster) populations to the emerging pitch canker pathogen, Fusarium circinatum. *PLoS ONE*, **9**, e114971.

FAO (2006) *Better forestry, less poverty; A practitioner's guide*, Food and Agriculture Organization of the United Nations, Rome.

Fernandes, P.M. & Rigolot, E. (2007) The fire ecology and management of maritime pine (*Pinus pinaster* Ait.). *Forest Ecology and Management*, **241**, 1–13.

Ganley, R.J., Watt, M.S., Manning, L. & Iturritxa, E. (2009) A global climatic risk assessment of pitch canker disease. *Canadian Journal of Forest Research*, **39**, 2246–2256.

García, J., De-Miguel, J., Alía, R. & S, I. (2001) *Regiones de Identificación y Utilización de material forestal de reproducción*, Ministerio de Medio Ambiente. Serie Cartográfica, Madrid.

Gaspar, M.J., Velasco, T., Feito, I., Alía, R. & Majada, J. (2013) Genetic variation of drought tolerance in *Pinus pinaster* at three hierarchical levels: A comparison of induced osmotic stress and field testing. *PLoS ONE*, **8**, 1–10.

Giljohann, K.M., Hauser, C.E., Williams, N.S.G. & Moore, J.L. (2011) Optimizing invasive species control across space: willow invasion management in the Australian Alps. *Journal of Applied Ecology*, **48**, 1286–1294.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distruction models in ecolgy. *Ecological Modelling*, **135**, 147–186.

Hair Jr, J.F., Anderson, R.E., Tatham, R.L. & William, C. (1995) *Multivariate data analysis*, Prentice Hall., New Jersey, USA.

Hamann, A. & Aitken, S.N. (2013) Conservation planning under climate change: Accounting for adaptive potential and migration capacity in species distribution models. *Diversity and Distributions*, **19**, 268–280.

Hampe, A. & Petit, R.J. (2005) Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, **8**, 461–467.

Hartkamp, A.D., Beurs, K. De, Stein, A. & White, J.W. (1999) Interpolation techniques for climate variables. *NRG-Geographic Information Systems*, **Series 99-**.

Hastie, T.J. & Tibshirani, R. (1990) *Generalized additive models*, Chapman & Hall/CRC, Boca Raton - London - New York - Washington D.C.

Heikkinen, R.K., Luoto, M., Virkkala, R., Pearson, R.G. & Körber, J.H. (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, **16**, 754–763.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hutchinson, M.F. (1991) *The application of thin plate smoothing splines to continent-wide data assimilation. Data Assimilation Systems*, (ed. by J.D. Jasper) BMRC Research Report No. 27, Bureau of Meteorology, Melbourne, Australia.

IPCC (2013) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, (ed. by T.F. Stocker), D. Qin), G.K. Plattner), M. Tignor), S.K. Allen), J. Boschung), A. Nauels), Y. Xia), V. Bex), and P.M. Midgley) Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,.

Iturritxa, E., Ganley, R.J., Raposo, R., García-Serna, I., Mesanza, N., Kirkpatrick, S.C. & Gordon, T.R. (2013) Resistance levels of Spanish conifers against Fusarium circinatum and Diplodia pinea. *Forest Pathology*, **43**, 488–495.

Iturritxa, E., Ganley, R.J., Wright, J., Heppe, E., Steenkamp, E.T., Gordon, T.R. & Wingfield, M.J. (2011) A genetically homogenous population of *Fusarium circinatum* causes pitch canker of *Pinus radiata* in the Basque Country, Spain. *Fungal Biology*, **115**, 288–295.

Iturritxa, E., Mesanza, N. & Brenning, A. (2015) Spatial analysis of the risk of major forest diseases in Monterey pine plantations. *Plant pathology*, **64**, 880–889.

Iturritxa, E., Mesanza, N., Elvira-Recuenco, M. Serrano, Y., Quintana, E. & Raposo, R. (2012) Evaluation of genetic resistance in Pinus to pitch canker in Spain. *Australasian Plant Pathology*, **41**, 601–607.

Jarvis, C.H. & Stuart, N. (2001) A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part I: The selection of “guiding” topographic and land cover variables. *Journal of Applied Meteorology*, **40**, 1060–1074.

Johnston, M.H., Campagna, M., Gray, P.A., Kope, H.H., Loo, J.A., Ogden, A.E., O’Neill, G.A., Price, D.T., Williamson, T.B. onM. H., Campagna, M., Gray, P.A., Kope, H.H., Loo, J.A., Ogden, A.E., O’Neill, G.A., Price, D.T. & Williamson, T.B. (2009) *Vulnerability of Canada’s tree species to climate change and management options for adaptation: an overview for policy makers and practitioners.*, Canadian Council of Forest Ministers, Ottawa, Ontario.

Jung, T. (2009) Beech decline in Central Europe driven by the interaction between *Phytophthora* infections and climatic extremes. *Forest pathology*, **39**, 73–94.

Landeras, E., García, P., Fernández, Y., Braña, M., Fernández-Alonso, O., Méndez-Lodos, S., Pérez-Sierra, A., León, M. & Abad-Campos, P. Berbegal, M.; Beltrán, R.; García-Jiménez, J.; Armengol, J. (2005) Outbreak of pitch canker caused by *Fusarium circinatum* on *Pinus* spp. in Northern Spain. *Plant Disease*, **89**, 1015–1015.

Madrigal-González, J. & Zavala, M.A. (2014) Competition and tree age modulated last century pine growth responses to high frequency of dry years in a water limited forest ecosystem. *Agricultural and Forest Meteorology*, **192-193**, 18–26.

Martinez-Alvarez, P., Pando, V. & Diez, J. (2014) Alternative species to replace Monterey pine plantations affected by pitch canker caused by *Fusarium circinatum* in northern Spain. *Plant Pathology*, **63**, 1086–1094.

McCullagh, P. & Nelder, J.A. (1989) *Generalized linear models*, (ed. by C. and Hall).

Monjo, R., Caselles, V. & Chust, G. (2014) Probabilistic correction of RCM precipitation in the Basque Country (Northern Spain). *Theoretical and Applied Climatology*, **117**, 317–329.

Morin, X. & Thuiller, W. (2009) Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology*, **90**, 1301–13.

Möykkynen, T., Capretti, P. & Pukkala, T. (2014) Modelling the potential spread of *Fusarium circinatum*, the causal agent of pitch canker in Europe. *Annals of Forest Science*, **72**, 169–181.

New, M., Hulme, M. & Jones, P.D. (1999) Representing twentieth century space-time climate variability. Part 1: development of a 1961-90 mean monthly terrestrial climatology. *Journal of Climate*, **12**, 829–856.

Normand, S., Ricklefs, R.E., Skov, F., Bladt, J., Tackenberg, O. & Svenning, J.C. (2011) Postglacial migration supplements climate in determining plant species ranges in Europe. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 3644–3653.

Pautasso, M., Dehnen-Schmutz, K., Holdenrieder, O., Pietravalle, S., Salama, N., Jeger, M.J., Lange, E. & Hehl-Lange, S. (2010) Plant health and global change - Some implications for landscape management. *Biological Reviews*, **85**, 729–755.

Pautasso, M., Schlegel, M. & Holdenrieder, O. (2015) Forest health in a changing world. *Microbial Ecology*, **69**, 826–842.

Pearson, R.G. (2010) Species' distribution modeling for conservation educators and practitioners. *Lessons in Conservation*, **3**, 54–89.

Pérez-Sierra, A., Landeras, E., León, M., Berbegal, M., García-Jiménez, J. & Armengol, J. (2007) Characterization of *Fusarium circinatum* from *Pinus* spp. in northern Spain. *Mycological research*, **111**, 832–839.

Preston, K.L., Rotenberry, J.T., Redak, R.A. & Allen, M.F. (2008) Habitat shifts of endangered species under altered climate conditions: importance of biotic interactions. *Global change biology.*, **14**, 2501–2515.

Prieto-Recio, C., Martín-García, J., Bravo, F. & Diez, J.J. (2015) Unravelling the associations between climate, soil properties and forest management in *Pinus pinaster* decline in the Iberian Peninsula. *Forest Ecology and Management*, **356**, 74–83.

R Development Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing,.

Redondo, M.A., Boberg, J. & Olsson, C.H.B. (2015) Winter conditions correlate with *Phytophthora alni* subspecies distribution in Southern Sweden. *Phytopathology*, **draft**, 1–25.

Ribalaygua, J., Torres, L., Pórtoles, J. avier, R, M., Gaitán, E. & Pino, M.R. (2013) Description and validation of a two-step analogue/regression downscaling method. *Theoretical and Applied Climatology*, **111**.

Rosso, P.H. & Hansen, E.M. (2003) Predicting swiss needle cast disease distribution and severity in young douglas-fir plantations in coastal Oregon. *Phytopathology*, **93**, 790–798.

Sakamoto, J.M. & Gordon, T.R. (2006) Factors influencing infection of mechanical wounds by *Fusarium circinatum* on Monterey pines (*Pinus radiata*). *Plant Pathology*, **55**, 130–136.

Sakamoto, J.M., Gordon, T.R., Storer, A.J. & Wood, D.L. (2007) The role of *Pityophthorus* spp. as vectors of pitch canker affecting *Pinus radiata*. *The Canadian Entomologist*, **139**, 864–871.

Santini, A., Ghelardini, L., De Pace, C., Desprez-Loustau, M.L., Capretti, P., Chandelier, A., Cech, T., Chira, D., Diamandis, S., Gaitniekis, T., Hantula, J., Holdenrieder, O., Jankovsky, L., Jung, T., Jurc, D., Kirisits, T., Kunca, A., Lygis, V., Malecka, M., Marcais, B., Schmitz, S., Schumacher, J., Solheim, H., Solla, A., Szabò, I., Tsopelas, P., Vannini, A., Vettraino, a. M., Webber, J., Woodward, S. & Stenlid, J. (2013) Biogeographical patterns and determinants of invasion by forest pathogens in Europe. *New Phytologist*, **197**, 238–250.

Schueler, S., Falk, W., Koskela, J., Lefèvre, F., Bozzano, M., Hubert, J., Kraigher, H., Longauer, R. & Olrik, D.C. (2014) Vulnerability of dynamic genetic conservation units of forest trees in Europe to climate change. *Global Change Biology*, **20**, 1498–1511.

Serra-Varela, M.J., Alía, R., Ruíz-Daniels, R., Zimmermann, N.E., Gonzalo-Jiménez, J. & Grivet, D. (in revision) Assessing vulnerability of two Mediterranean conifers to support genetic conservation management. *Diversity and Distributions*.

Serra-Varela, M.J., Grivet, D., Vincenot, L., Broennimann, O., Gonzalo-Jiménez, J. & Zimmermann, N.E. (2015) Does phylogeographic structure relate to climatic niche divergence? A test using maritime pine (*Pinus pinaster* Ait.). *Global Ecology and Biogeography*, **24**, 1302–1313.

Sturrock, R.N., Frankel, S.J., Brown, A. V., Hennon, P.E., Kliejunas, J.T., Lewis, K.J., Worrall, J.J. & Woods, A.J. (2011) Climate change and forest diseases. *Plant Pathology*, **60**, 133–149.

Taylor, K.E., Stouffer, R.J. & Meehl, G.A. (2009) A Summary of the CMIP5 Experiment Design. *World*, **4**, 1–33.

Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009) BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369–373.

Thuiller, W., Richardson, D.M., Pyšek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, **11**, 2234–2250.

Tylianakis, J.M., Didham, R.K., Bascompte, J. & Wardle, D.A. (2008) Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, **11**, 1351–1363.

Valladares, F., Matesanz, S., Guilhaumon, F., Araújo, M.B., Balaguer, L., Benito-Garzón, M., Cornwell, W., Gianoli, E., van Kleunen, M., Naya, D.E., Nicotra, A.B., Poorter, H. & Zavala, M. a. (2014) The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecology Letters*, **17**, 1351-1364.

Vivas, M., Zas, R. & Solla, A. (2012) Screening of maritime pine (*Pinus pinaster*) for resistance to *Fusarium circinatum*, the causal agent of pitch canker disease. *Forestry*, **85**, 185–192.

Walther, G., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J., Hoegh-Guldberg, O. & Bairlein, F. (2012) Ecological responses to recent climate change. *Nature*, **416**, 389–395.

Watts, M.J., Ganley, R.J., Kriticos, D.J. & Manning, L.K. (2011) Dothistroma needle blight and pitch canker: the current and future potential distribution of two important diseases of *Pinus* species. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, **41**, 412–424.

Welsh, C., Lewis, K.J. & Woods, A.J. (2014) Regional outbreak dynamics of Dothistroma needle blight linked to weather patterns in British Columbia, Canada. *Canadian Journal of Forest Research*, **44**, 212–219.

Zas, R., Sampedro, L., Prada, E. & Fernández-López, J. (2005) Genetic variation of *Pinus pinaster* Ait. seedlings in susceptibility to the pine weevil *Hylobius abietis* L. *Annals of Forest Science*, **62**, 681–688.

Zorita, E. & Von Storch, H. (1999) The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate*, **12**, 2474–2489.

Supporting Information

Additional Supporting Information may be found in the annex section.

Appendix S1. Details on meteorological stations and climatic data

Fig S1: Geographical representation and histogram of elevations of the meteorological stations from AEMET and WORLDCLIM

Fig S2: Selected environmenal variables for *Pinus pinaster* Ait. (a) and pitch canker disease (b) to be included within their species distribution models.

Chapter 4

Integrating geography and environment in a standardized procedure to assess marginality and its effects on genetic patterns. An example with European conifers.

Integrating geography and environment in a standardized procedure to assess marginality and its effects on genetic patterns. An example with European conifers.

Serra-Varela *et al.* in prep.

Authors (to be confirmed; in alphabetical order)

Serra-Varela, M.J.[1,2,3]	maje.serra@gmail.com
Alía, R. [2,3]	alia@inia.es
Aravanopoulos, P.	aravanop@for.auth.gr
Cavers, S.	scav@ceh.ac.uk
Fady, B.	bruno.fady@avignon.inra.fr
Gonzalo-Jiménez, J. [2,3]	jgonzalo@pvs.uva.es
González-Martínez, S.C.	santiago.c.gonzalez.martinez@gmail.com
Grivet, D. [2,3]	dgrivet@inia.es
Marchi, M.	maurizio.marchi@crea.gov.it
Notivol, E.	enotivol@aragon.es
Picard, N.	nicolas.picard@fao.org
Piotti, A.	andrea.piotti@gmail.com
Vendramin, G.	gg.vendramin@gmail.com
Westergren, M.	marjana.westergren@gozdis.si

Abstract

Genetic patterns in marginal populations have been largely studied, as these populations may harbour unique genetic combinations due to their location in the rear/leading edge of species distribution. However, according to the abundant-centre hypothesis, they may have low genetic diversity due to patchy distributions and small population sizes. Until the date, empirical studies are controversial, the main reason being that we lack standardized definition of marginality, which would require considering processes affecting gene flow, such as isolation by distance (IBD) and isolation by environment (IBE). Here, we provide a consistent approach to define marginality integrating each of these two processes, by separately accounting for geographic and environmental factors respectively. We based our methodology on a set of indices (both geography and environment related) that can be directly integrated, in a subsequent step, with genetic diversity indices, to further understand the relationship among these three elements. We illustrated our work with four Mediterranean conifers, namely, *Pinus halepensis* Mill., *Pinus pinaster* Ait., *Pinus nigra* Arnold. and *Pinus pinea L.*, as well as three European conifers namely *Abies alba* Mill., *Pinus sylvestris* L. and *Picea abies* L. These broadly-distributed European species are suitable for our analysis, as they have been genetically characterized, enabling the development of the second stage of the work (*i.e.* the effect of marginality on genetic patterns).

We approached geographic marginality by means of four geographic indices, namely, core size, distance to the border of the core, cost-distances to the centroid of the distribution and distance to the nearest large patch, aiming to characterize the position of each population within its core patch and in relation to the entire distribution of the species. We assessed environmental marginality by fitting individual species distribution models (SDMs) for each species, characterizing each location with a probability of environmental suitability (environmental index). For each species, we considered as environmentally marginal, those populations whose environmental indices were below the 5th and the 10th percentile of its entire distribution. Finally, we calculated the above described indices for all genetically characterized populations.

Because this study is still undergoing at the time this manuscript is being produced, we will present here the concepts and methodology related to the current available results, as well as comments on the preliminary results. We expect to have a definitive report by the end of 2016.

Keywords: Species Distribution Model (SDMs), Morphological Spatial Pattern Analysis (MSPA), Isolation by Distance (IBD), Isolation by Environment (IBE).

1. Introduction

Marginal or peripheral populations, have been a focus of interest since the 70's (see Soule, 1973), aiming to disentangle the effect of marginality on genetic diversity and on fitness related traits, and resulting in the development of different theoretical hypotheses. Genetic variation in marginal populations could be shaped by past climate, and these populations may therefore accumulate important and unique genetic combinations pivotal to the species' adaptation to climate change (Hardie & Hutchings, 2010), particularly at the edges of the distribution (Hampe & Petit, 2005). Thereby, although we should not presuppose their importance (Lesica & Allendorf, 1992; Hardie & Hutchings, 2010), their implications for conservation management must be carefully analysed. According to the abundant-centre hypothesis (Hengeveld & Haeck, 1982), these populations are likely to be smaller and more spatially isolated than in the core, where species are supposed to achieve their highest abundance. Thus, marginal populations are expected to have lower genetic diversity due to genetic drift and inbreeding (Frankham, 1996), as well as lower fitness than core populations, given that local adaptation could be constrained by incoming migrants from more central populations. However, gene flow from central populations could also increase genetic diversity and consequently the probabilities of a successful adaptation (see Kremer *et al.*, 2012 for a review). Empirical studies have not been able to solidly support theoretical assumptions as they have reported contradictory results, reviewed in Eckert *et al.* (2008) and Sagarin & Gaines (2002), the main reason being that they differ in the use of the term marginality. Consequently, there is an urgent necessity to develop a standardized and objective procedure to define marginal populations, in the context of meeting theoretical predictions and empirical results.

The effects of marginality on genetic patterns are driven by the same processes that underlie species' genetic structure *i.e.* gene flow and genetic drift (Orsini *et al.*, 2013). While the latter is mainly affected by population size, the former is influenced by different factors within which geography and environment play a major role. Geographic distance constrains gene flow and thus we expect nearer populations to be more genetically similar than more separated ones (Isolation by distance pattern; IBD - Wright, 1943). Similarly, in the case of isolation by environment (IBE - Wang & Bradburd, 2014), genetic resemblance is associated to environmental similarity, as gene flow is more likely to be successful between similar environments (Sexton *et al.*, 2014). Thus, when assessing the effects of marginality on genetic patterns, both processes should be considered. In fact, both are contemplated within the framework of the abundant-centre hypothesis, although this theory assumes that geography and environment vary together and consequently, populations inhabiting peripheral locations experience extreme environmental conditions (compared to the rest of the distribution) close to the limit of species' physiology. However, in practice, it is common to find peripheral populations inhabiting environmental conditions similar to those of the core, as well as the opposite scenario where core populations occupy extreme habitats: these situations claim for the need to consider separately both concepts of marginality, *i.e.* geographic marginality *vs* environmental marginality. Indeed, despite that geography is the most commonly used factor to define marginal populations (see Jiménez *et al.*, 1999; Hamrick *et al.*, 1989; Yeh & Layton, 1979), topographic (see Angert, 2006 for an example with altitude) and even environmental variables (see Lira-Noriega & Manthey, 2014; Diniz-Filho *et al.*, 2009) have already been employed to test the abundant-centre hypothesis.

Here, we provide a framework to assess both concepts of marginality (geographical and environmental) following an objective and standardized procedure in order to, in a second stage, assess its consequences on genetic diversity. We approach geographic marginality by means of four indices relevant from a genetic point of view. Two of them, core size and distance to the border of the core, represent the location of the population within its core patch, while the other two, namely, cost-distance to the centroid of the distribution and distance to a large core patch, illustrate the position of the population with respect to the entire distribution of the species. We approach environmental marginality by means of species distribution models (SDMs; see Guisan & Zimmermann, 2000 for details) that associate a probability value to each population representing its habitat suitability.

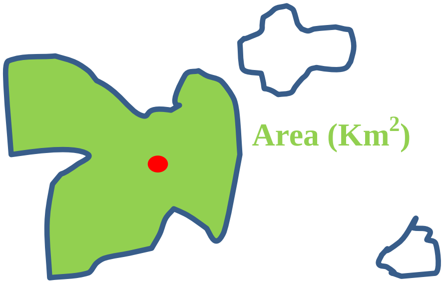
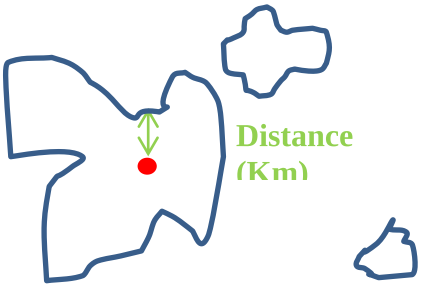
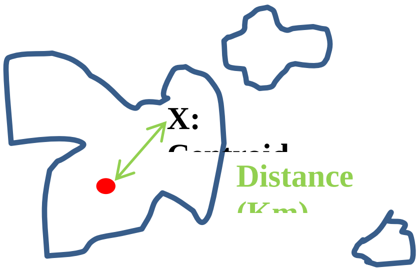
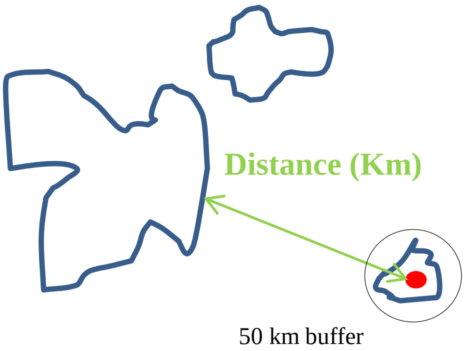
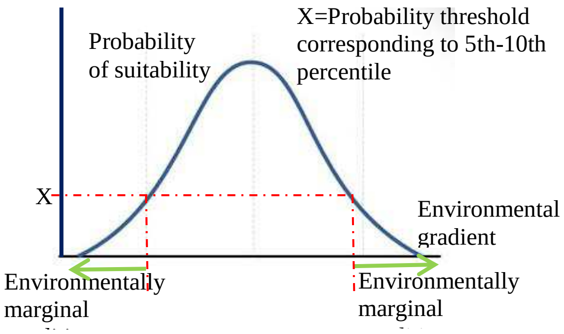
We exemplify our approach with seven widely distributed European conifers namely, *Pinus halepensis* Mill., *Pinus pinaster* Ait., *Pinus nigra* Arnold., *Pinus pinea L., Abies alba* Mill., *Pinus sylvestris* L. and *Picea abies* L. including (but for two exceptions) their entire distribution ranges. These species are of large interest for the scientific community, and as such they are genetically characterized for a set of populations, based on different genetic markers available – namely, isozymes, single nucleotide polymorphisms (SNPs) and nuclear or cloroplast simple sample repeats (cp/n-SSR). This available information will allow performing the second stage of the analysis, *i.e.* studying the effect of marginality on genetic patterns. Moreover, developing a multi-specific approach will enable to analyse whether there are marginality hotspots encompassing different species' marginal populations.

2. Materials and Methods

For each species, and on the basis of its distribution and its genetically characterized populations, we calculated five indices to assess geographic and environmental marginality (see Table 1). Geographical indices were based exclusively on geometrical measures, aiming to summarize those characteristics of the distribution that are likely to affect gene flow. To assess environmental marginality we used SDMs, based on the distribution of species and climatic variables from WORLDCLIM (Hijmans *et al.*, 2005). These models enabled to characterize each location with a probability of suitability (here on environmental index), based on its climatic conditions and to assess the percentile they occupy with respect to the entire distribution.

We assessed the distribution of the seven species based on EUFORGEN polygons (http://www.euforgen.org/distribution_maps.html; which map exclusively native populations), filtered with the more spatially accurate Tree Species Distribution for Europe (TSDE; Köble & Seufert, 2001) from the Joint Research Center's AFOLU data portal (ftp://mars.jrc.ec.europa.eu/Afoludata/Public/DS66/).

Table 1: Indexes employed to assess geographical and environmental marginality. Blue polygons represent the distribution. Red dots are genetically characterized populations.

Geographic indices	Position within the core	Size of the core patch	
		Distance to the border of the core	
	Position within the distribution	Cost-distance to the centroid of the distribution	
		Distance to the nearest large patch outside a 50 Km buffer	
Environmental index	Probability of habitat suitability	Environmental marginality: populations whose environmental index is below the 5th or 10th percentile	

2.1. Geographic marginality

2.1.1. Morphological Spatial Pattern Analysis - MSPA

First, we smoothed distribution's noisy boundaries, resulting from the combination of EUFORGEN and TSDE, as they impeded a correct performance of the analysis. We addressed this issue by performing an erosion/dilation process using a 2x2 kernel in the "EBImage" package in R statistical environment (R Development Core Team, 2016). Subsequently, and using the same package, we performed a morphological spatial pattern analysis (MSPA - Soille & Vogt, 2008), a customized sequence of mathematical morphological operators targeted at the description of the geometry and connectivity of the components of a binary image. In our particular case, the binary image corresponded to the species' distribution and we differentiated foreground data (presence cells), background data (absence cells) and non-available data (in this case corresponding to the sea cells). Then, by means of our MSPA, exclusively based on geometric concepts, we divided the distribution into three generic MSPA classes: i) cores *i.e.* continuous patches of presences excluding their perimeters, ii) edges, which corresponded to external cores' perimeters (1 Km width), and iii) others, that ingathered other categories such as islets, branches or loops as we applied a simplified version of the categories proposed in the JRC Guidos Tool Box (see Vogt, 2016).

2.1.2. Geographic marginality indices

We calculated geographic indices for genetically characterized populations. Geographic indices aimed to characterize both, (i) the position of the population within its corresponding core, (ii) the position of the population in relation to the entire distribution of the species.

Position within the patch was characterized by means of (i) core size and (ii) distance to the edge of the core. When target populations were located outside a core, we calculated the size of their nearest core, and associated a negative sign to its distance to the nearest edge. To assess the position of populations in relation to the rest of the distribution, we estimated two additional indices. First, we calculated least cost-distances to the centroid of the distribution. We used a least cost patch analysis in order to avoid non-possible paths. For this aim, we constructed a gene flow conductance matrix (the opposite of a resistance matrix) considering the sea to have low conductance (1), the land as having an intermediate conductance (50) and other populations of the species to be highly conducting (100). We performed this analysis by means of the “gdistance” package in R. Second, we aimed to assess the level of broad-scale connectivity of populations. To do so, we assessed their proximity (distance) to the nearest large core - larger than 100 ha, as this is our resolution size – and outside a 50 Km buffer - to prevent the consideration of very close cores.

2.2. Environmental marginality

2.2.1. Occurrence records

Our study area, encompassing the European continent and northern Africa, covers the entire distribution of all target species but for *P. sylvestris* and *Picea abies*, which spread largely into the Asian continent. We set the eastern limits of our study area to the Ural Mountains, as this orographic feature constitutes an important barrier to gene flow, resulting in different subspecies of *P. sylvestris* and *P. abies* inhabiting each slope. Thereby, for these two species, SDMs are exclusively based on European occurrences, as these models have been reported to better perform when fitted to sub-species rather than to the species level (see Serra-Varela *et al.*, 2015; D’Amen *et al.*, 2013 for examples).

Presence records were those included simultaneously in EUFORGEN and TSDE. Possible pseudo-absences corresponded to all the rest of the territory within the study area where TSDE reported 0 % occupancy. The numbers of presences within our defined area amounted to: 135,737 for *P. halepensis*, 128,744 for *P. pinaster*, 363,620 for *P. nigra*, 30,591 for *P. pinea*, 8,222,990 for *P. sylvestris*, 247,966 for *A. alba* and 7,667,354 for *P. abies*. The number of pseudo-absences and the method to select them is specified below together with algorithm selection.

2.2.2. Bioclimatic data

We downloaded the 19 bioclimatic variables available in WORLDCLIM (Hijmans *et al.*, 2005) representative of the period 1950-2000 for the analysis and selected for each species, four or five variables, relevant and low correlated variables, both temperature- and precipitation-related, to characterize their bioclimatic niche (see Table 2). For more details on variable selection see Appendix S1.

Table 2: Set of relevant and weakly correlated bioclimatic predictors selected for each target species

SPECIES	BIOCLIMATIC PREDICTORS				
Pinus halepensis *Pinus pinaster*	BIO1 (Annual Mean Temperature)	BIO4 (Temperature Seasonality)	BIO18 (Precipitation of Warmest Quarter)	BIO19 (Precipitation of Coldest Quarter)	-
Pinus nigra	BIO2 (Mean diurnal range)	BIO4 (Temperature Seasonality)	BIO11 (Mean Temperature of Coldest Quarter)	BIO19 (Precipitation of Coldest Quarter)	-
Pinus pinea	BIO1 (Annual Mean Temperature)	BIO4 (Temperature Seasonality)	BIO15 (Precipitation seasonality)	BIO18 (Precipitation of Warmest Quarter)	BIO19 (Precipitation od Coldest Quarter)
Abies alba	BIO3 (Isothermality)	BIO7 (Temperature Annual Range)	BIO10 (Mean Temperature of Warmest Quarter)	BIO16 (Precipitation of Wettest Quarter)	BIO18 (Precipitation of Warmest Quarter)
Pinus sylvestris *Picea abies*	BIO3 (Isothermality)	BIO4 (Temperature Seasonality)	BIO10 (Mean Temperature of Warmest Quarter)	BIO15 (Precipitation seasonality)	BIO18 (Precipitation of Warmest Quarter)

2.2.3 Species Distribution Models

We used Generalized Additive Models (GAM; Hastie & Tibshirani, 1990) to model the distribution of the target species, processed in BIOMOD (Thuiller *et al.*, 2009) using the package "biomod2" (default settings for the model) in the R statistical software environment (R Development Core Team, 2016). We randomly selected a large number of pseudoabsences (five times the number of presences except for *P. sylvestris* - 16,426,602 - and *P. abies* – 16,620,411 - as their large extent, and thus large number of presence records, resulted in a too large number of required absences for which there were not enough cells within the study area), and the same weight was given to presences and absences as recommended for GAM by Barbet-Massin *et al.*, (2012). Model performance was assessed by means of True Skill Statistic (TSS; Allouche *et al.*, 2006) and Area Under the ROC Curve (AUC; Fielding

& Bell, 1997). The large number of occurrence records available permitted a random division of each dataset into two equally-sized subsets for training and evaluating. Both subsets maintained the initial proportion between presence and pseudo-absence records. Probabilistic model outputs based on current climatic conditions were converted to binary maps (environmentally suitable *vs* environmentally unsuitable) by defining thresholds that optimized TSS values.

2.2.4 Environmental marginality: Locations and indices

Based on the SDM-probabilistic projections, we calculated environmental indices for genetically characterized populations as the probability of suitability of their locations. This way, we prepared the second-stage analysis of this work in which we will integrate environmental and geographical marginality to assess their effect on genetic diversity. In addition, we assessed the percentile occupied by the environmental index of each population considering the entire distribution of the species.

Environmentally marginal populations are those inhabiting rare or extreme environmental conditions compared with the rest of the species' distribution. Accordingly, to detect them, we selected those populations whose environmental indices (*i.e.* probabilities of environmental suitability) were below the fifth and tenth percentile within the entire distribution of the species. Then, by overlapping in the geographic space the environmentally marginal populations detected (only for the 10th percentile as these included the 5^{th}-percentile) corresponding to the seven analysed species, we assessed possible environmentally marginal hotspots across our study area. To facilitate the interpretation of the results, we also obtained the level of overlap among the current distribution of all our target species.

3. Results and Discussion

3.1.- Geographic marginality

3.1.1. Morphological Spatial Pattern Analysis - MSPA

The MSPA enabled to disaggregate the distribution of species in cores, edges or others (see Fig. 1 for an example with *P. pinaster*). The number of cores detected, which depended on the extent of the species and on its fragmentation, varied among species as follows: *P. halepensis* (122 cores); *P. pinaster* (115 cores); *P. nigra* (70 cores); *P. pinea* (55 cores); *A. alba* (241 cores); *P. sylvestris* (1608 cores); *P. abies* (598 cores). The number of cores majorly depends on the extent of the distribution of the species. However, among species that occupy similar extensions, higher number of cores would indicate higher fragmentation.

3.1.2. Geographic marginality indices

The results for this section are still on going. Those already available can be checked in Tables S2-S7 in Appendix S1.

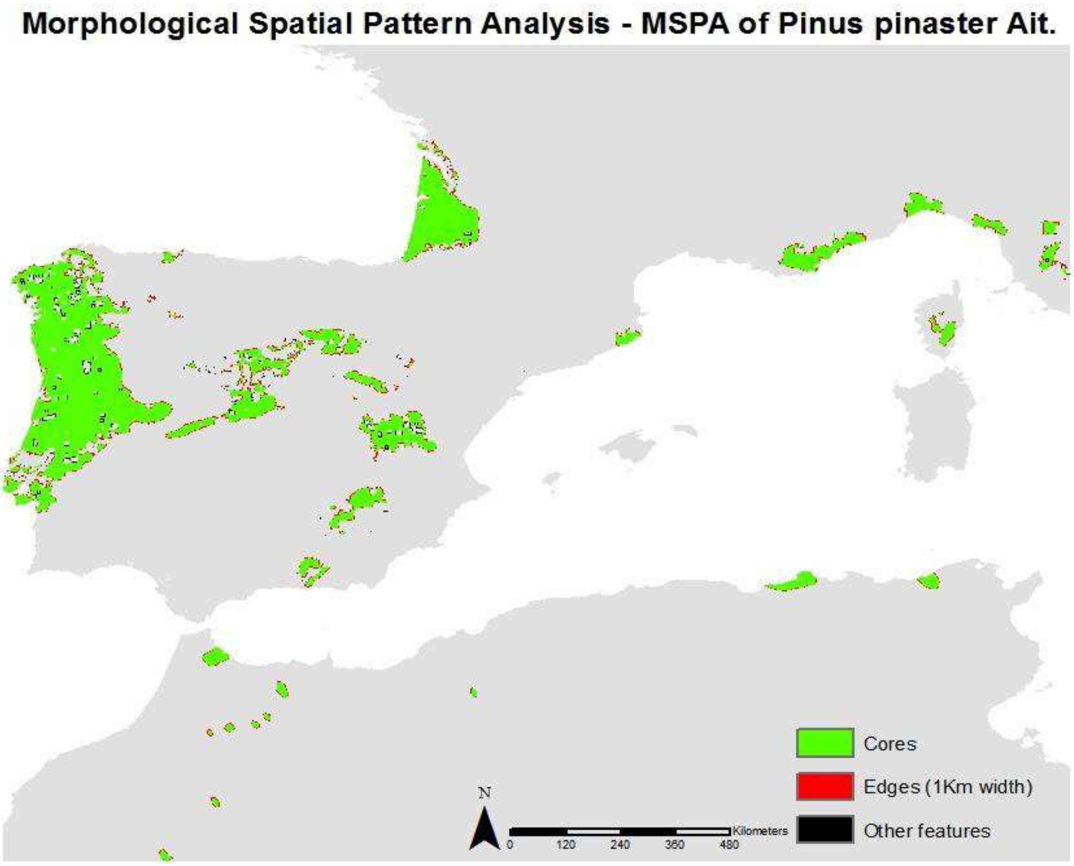


Figure1: Morphological Spatial Pattern Analysis – MSPA (Soille & Vogt, 2008), exemplified by *Pinus pinaster* Ait.

3.2. Environmental marginality

3.2.1 Species distribution models

Our models obtained moderate to high performance as revealed by AUC and TSS scores (Table 3). AUC scores were all very high (above 0.94 for all species), while TSS values showed a wider range of values (from 0.74 in *P. sylvestris* to 0.898 in *P. pinea*). Sensitivity and specificity scores were also elevated, although sensitivity displayed higher values than specificity in all cases. Thereby, models were better able to adequately identify suitable than unsuitable locations.

Table 3: Area Under the ROC Curve (AUC) and True Skill Statistic (TSS) scores from the Species Distribution Models fitted for the different species. Cut-off values (optimizing TSS) to convert probabilistic projections into binary models are also included as well as their associated Sensitivity and Specificity.

Species	AUC	TSS	Cutoff value (optimizing TSS)	Sensitivity (optimizing TSS)	Specificity (optimizing TSS)
P. halepensis	0.963	0.880	53.9	97.875	90.164
P. pinaster	0.976	0.876	56.5	96.311	91.245
P. nigra	0.962	0.805	59.3	92.492	88.055
P. pinea	0.984	0.898	55.6	95.921	93.836
A. alba	0.980	0.892	45.5	97.450	91.743
P. sylvestris	0.945	0.740	52.5	89.058	84.933
P. abies	0.977	0.845	46.5	95.096	89.387

Probabilistic and binary geographic projections for all species are shown in Appendix S1 (Fig. S1 - Fig. S7). In the purely Mediterranean species, *i.e.* *P. halepensis*, *P. pinaster*, *P. nigra* and *P. pinea*, the predicted projections (both probabilistic and binary) widely extended beyond the currently occupied territory of the species (Fig. S1 – Fig.S4), suggesting that there are additional reasons, further than climate, limiting these species' distributions (*e.g.* dispersal limitation, anthropogenic disturbances or pests and diseases), which was consistent with previous studies analyzing Mediterranean species (*e.g.* Serra-Varela *et al.*, 2015; in revision). On the contrary, the predicted projections for *P. abies* and *P. sylvestris* were very similar to their realized niches (*i.e.* current distribution) suggesting that, in this case, bioclimatic variables play a major role in driving these species' distribution.

3.2.2. Environmental marginality: Locations and indices

Environmentally marginal populations for our seven target species are shown in Appendix S1 (Fig. S1 – Fig. S7) and their corresponding environmental indices thresholds for the 5th and 10th percentiles are shown in Table 4. All species, but *P. pinea* (which threshold values were extremely low, see Table 4), showed relative high probability thresholds: all were above the TSS scores employed to separate suitable from unsuitable locations in binary projections (comparison of Tables 3 & 4), except for *P. sylvestris*. This result suggests that although some species inhabit very different habitats across their distributions (*e.g. Pinus pinaster* presents Atlantic and Mediterranean populations), their diverse habitats are represented by many populations, translating into few populations living in rare habitats.

Table 4: Environmental indices corresponding to the fifth and tenth percentiles used as thresholds to detect populations experiencing rare or extreme environmental conditions

Species	5th percentile	10th percentile
P. halepensis	66.7	77.8
P. pinaster	62.7	77.6
P. nigra	47.4	65.6
P. pinea	1.3	6.7
A. alba	58.9	76.7
P. sylvestris	28.1	50.1
P. abies	46.9	64.7

We found a maximum of four species coinciding in their distribution that concentrated in the Balkans and in the Spanish Mediterranean coast (Fig. 2a). Of these two regions, locations encompassing four-species environmentally marginal populations were very scarce and only found in the Balkans (see Fig. 2b). Majorly, we found one-species environmentally marginality locations, although scattered two-species locations could also be found across central Europe and the Baltic Peninsula (comparison between Fig. 2a & 2b).

We found a maximum of four species coinciding in their distribution that concentrated in the Balkans and in the Spanish Mediterranean coast (Fig. 2a). Of these two regions, locations encompassing four-species environmentally marginal populations were very scarce and only found in the Balkans (see Fig. 2b). Majorly, we found one-species environmentally marginality locations, although scattered two-species locations could also be found across central Europe and the Baltic Peninsula (comparison between Fig. 2a & 2b).

Environmental indices for genetically characterized populations are shown in Tables S2 - S7, except for the case of *P. nigra* for which the genetic information was still not available at the time we were preparing this document. *Abies alba*, *P. sylvestris*, and *P. abies* distinguished for the high number of populations characterized by environmental indices associated to low percentiles (see Fig. 3). *Pinus halepensis* and

P. pinaster showed a more homogeneous distribution across the different percentiles, while *P. pinea*, showed the contrary pattern with high percentiles being the most represented. These results indicated that for some species sampling was not homogenously performed across the environmental gradient occupied by the species. Thus, while *P. halepensis* and *P. pinaster* are well represented, the rest show a clear bias towards environmentally suitable locations (*P. pinea*) or environmentally unsuitable locations (*Abies alba*, *P. sylvestris*, and *P. abies*). Our analysis aims to understand the effects of marginality (both geographic and environmental) on genetic patterns, which would be more easily performed for those species (such as *P. pinaster* and *P. halepensis*) in which both types of populations (core and marginal) are represented. The unbalanced representation of populations concerns exclusively the environmental aspect of marginality, while a different pattern may be detected in the geographical aspect of marginality (ongoing analyses).

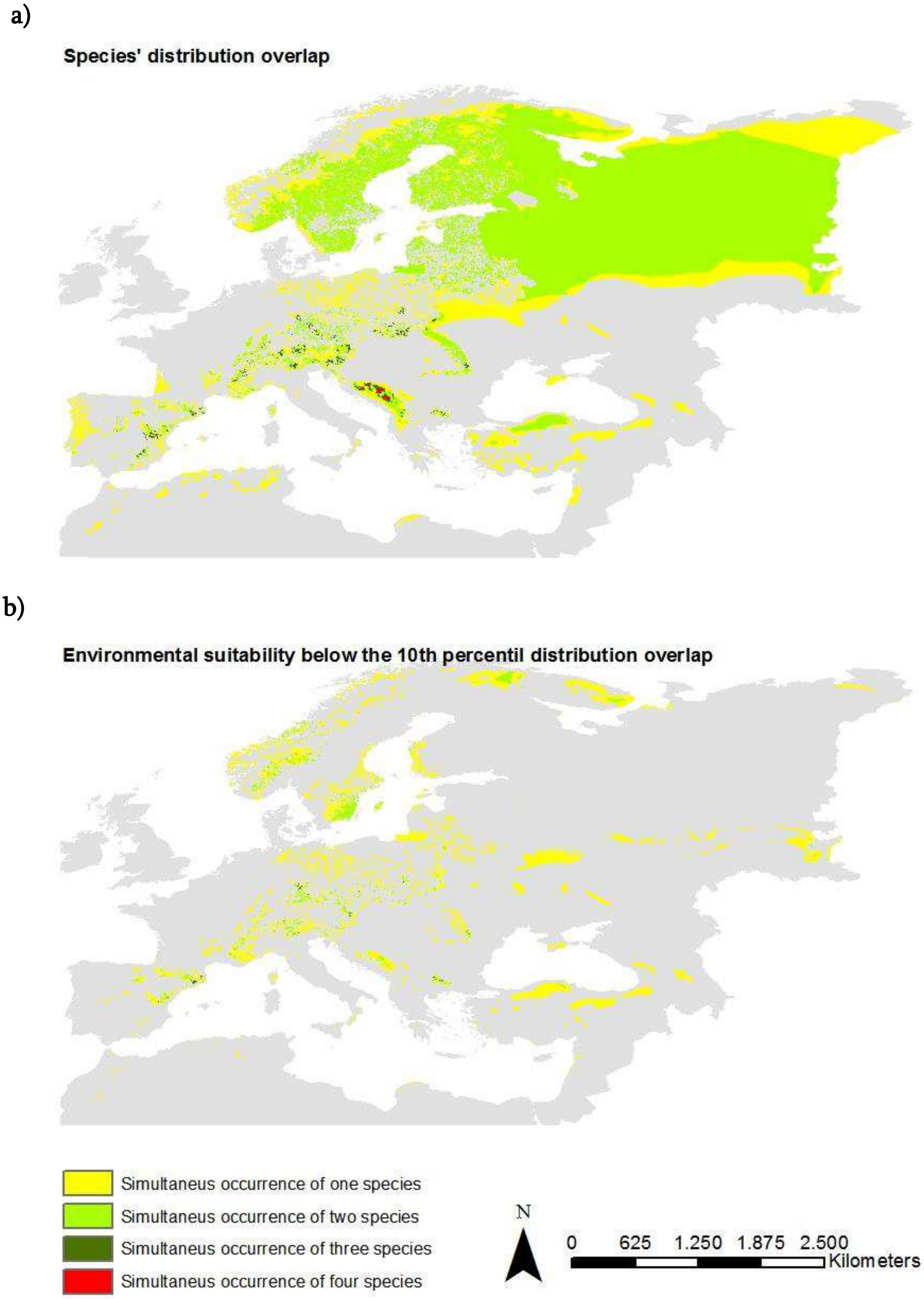
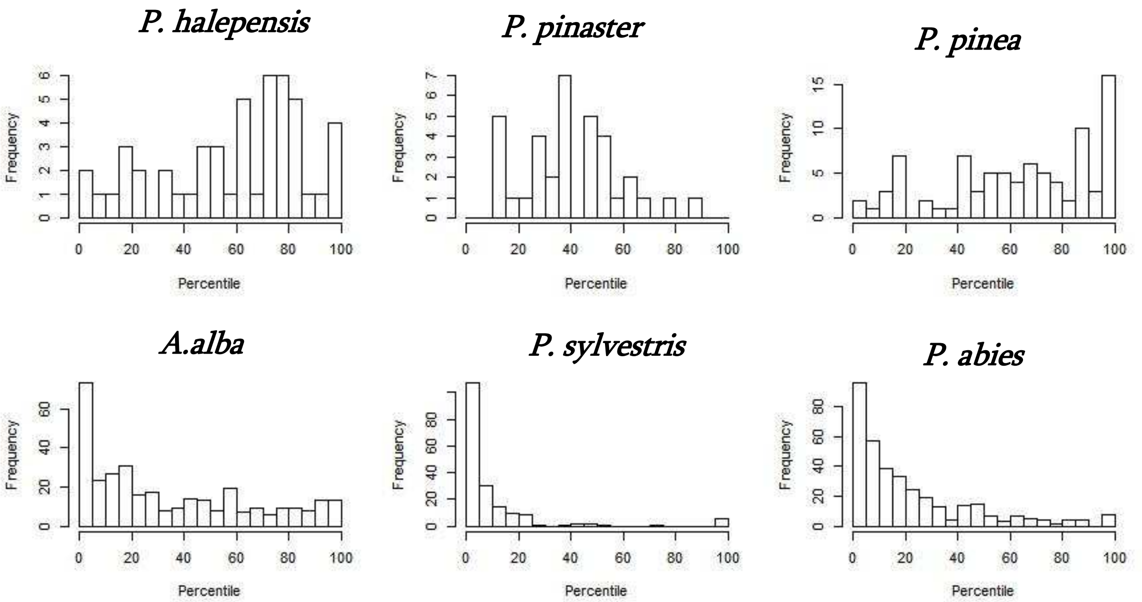


Figure 2 : Overlap among target species distribution (a) and distribution below the 10th environmental suitability percentile (b)



P. halepensis *P. pinaster* *P. pinea*

A.alba *P. sylvestris* *P. abies*

Figure 3 : Histograms of the percentiles from the genetically characterized populations for which we calculated environmental indexes (see Tables S.1-S.7 in Appendix S.1)

4. Conclusions

This chapter presents a novel and standardized methodology to assess marginality from two different perspectives, the environmental and geographical points of view. From our preliminary results on environmental marginality, we highlight: i) probability thresholds corresponding to the 10th percentile varied widely across species (from 6.7 in *P. pinea* to 77.8 in *P. halepensis*); ii) the very few overlap among species in terms of environmental suitability, presaged a reduced number of hotspot of environmental suitability across species; iii) it is important to select well the populations to be genetically characterized across all habitat types, to ensure a good representation of all environmental conditions. While the integration of geography to assess marginality is still undergoing, the final step of this study, will be to assess the effect of geographic and environmental marginality on genetic patterns. The fact that we will perform our analysis across a wide range of species, will contribute to ensure a solid evaluation of these effects.

Acknowledgements

We thank the COST ACTION FP 1202 that provided a framework for the development of this works and enabled the collaboration of the different groups involved.

References

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.

Angert, A.L. (2006) Demography of central and marginal populations of Monkeyflowers (*Mimulus cardinalis* and *M. Lewisii*). *Ecology*, **87**, 2014–2025.

Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **3**, 327–338.

D'Amen, M., Zimmermann, N.E. & Pearman, P.B. (2013) Conservation of phylogeographic lineages under climate change. *Global Ecology and Biogeography*, **22**, 93–104.

Diniz-Filho, J.A.F., Nabout, J.C., Bini, L.M., Soares, T.N., Pires de Campus Telles, M., de Marcon Jr, P. & Collevatti, R.G. (2009) Niche modelling and landscape genetics of *Caryocar brasiliense* ("Pequi" tree: Caryocaraceae) in Brazilian Cerrado: An integrative approach for evaluating central-peripheral population patterns. *Tree Genetics and Genomes*, **5**, 617–627.

Eckert, C.G., Samis, K.E. & Lougheed, S.C. (2008) Genetic variation across species' geographical ranges: The central-marginal hypothesis and beyond. *Molecular Ecology*, **17**, 1170–1188.

Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.

Frankham, R. (1996) Relationship of genetic variation to population size in wildlife. *Conservation Biology*, **10**, 1500–1508.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distruction models in ecology. *Ecological Modelling*, **135**, 147–186.

Hampe, A. & Petit, R.J. (2005) Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, **8**, 461–467.

Hamrick, J.L., Blanton, H.M. & Hamrick, K.J. (1989) Genetic structure of geographically marginal populations of ponderosa pine. *American Journal of Botany*, 1559–1568.

Hardie, D.C. & Hutchings, J.A. (2010) Genetic structure of geographically marginal populations of ponderosa pine. *Environmental Reviews*, **18**, 1–20.

Hastie, T.J. & Tibshirani, R. (1990) *Generalized additive models*, Chapman & Hall/CRC, Boca Raton - London - New York - Washington D.C.

Hengeveld, R. & Haeck, J. (1982) The Distribution of abundance. I. Measurements. *Journal* *of Biogeography*, **9**, 303–316.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Jiménez, P., Agúndez, D., Alía, R. & Gil, L. (1999) Genetic variation in central and marginal populations of *Quercus suber* L. *Silvae Genetica*, **48**, 278–284.

Köble, R. & Seufert, G. (2001) *Novel maps for forest tree species in Europe. Proceedings of the 8th European Symposium on the Physico-Chemical Behaviour of Air Pollutants: "A Changing Atmosphere!,"* pp. 17–20. Data set available online at: http://afoludata.jrc.ec.europa.eu/index.php/dataset/detail/66, Torino.

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

Lesica, P. & Allendorf, F.W. (1992) Are small populations of plants worth preserving? *Conservation Biology*, **6**, 135–139.

Lira-Noriega, A. & Manthey, J.D. (2014) Relationship of genetic diversity and niche centrality: A survey and analysis. *Evolution*, 1–43.

Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J. & De Meester, L. (2013) Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular ecology*, **22**, 5983–99.

R Development Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing,.

Sagarin, R.D. & Gaines, S.D. (2002) The "abundant centre" distribution: To what extent is it a biogeographical rule? *Ecology Letters*, **5**, 137–147.

Serra-Varela, M.J., Alía, R., Ruíz-Daniels, R., Zimmermann, N.E., Gonzalo-Jiménez, J. & Grivet, D. (in revision). Assessing vulnerability of two Mediterranean conifers to support genetic conservation management. *Diversity and Distributions*.

Serra-Varela, M.J., Grivet, D., Vincenot, L., Broennimann, O., Gonzalo-Jiménez, J. & Zimmermann, N.E. (2015) Does phylogeographic structure relate to climatic niche divergence? A test using maritime pine (*Pinus pinaster* Ait.). *Global Ecology and Biogeography*, **24**, 1302–1313.

Sexton, J.P., Hangartner, S.B. & Hoffmann, A.A. (2014) Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution*, **68**, 1–15.

Soille, P. & Vogt, P. (2008) Morphological segmentation of binary patterns. *Pattern Recognition Letters*, **30**, 456–459.

Soule, M. (1973) The epistasis cycle: A theory of marginal populations. *Annual Review of*

Ecology and Systematics, **4**, 165–187.

Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009) BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369–373.

Vogt, P. (2016) Guidos Tool Box (Graphical User Interface for the Description of image Objects and their Shapes): Digital image analysis software collection available at the following web site: http://forest.jrc.ec.europa.eu/download/software/guidos.

Wang, I.J. & Bradburd, G.S. (2014) Isolation by environment. *Molecular Ecology*, **23**, 5649–5662.

Wright, S. (1943) Isolation by Distance. *Genetics*, **28**, 114–138.

Yeh, F.C. & Layton, C. (1979) The organization of genetic variability in central and marginal populations of lodgepine *Pinus contorta spp. latifolia*. *Canadian Journal of Genetics and Cytology*, **21**, 487–503.

Supporting Information

Additional Supporting Information may be found in the annex section.

Appendix S1. Details on meteorological stations and climatic data

Fig S1: Geographical representation and histogram of elevations of the meteorological stations from AEMET and WORLDCLIM

Fig S2: Selected environmenal variables for *Pinus pinaster* Ait. (a) and pitch canker disease (b) to be included within their species distribution models.

Future perspectives

Future perspectives

This thesis provides several examples of the important role that SDMs can play in forest conservation management (Chapters 2, 3 & 4) while addressing three possible directions to improve their predictions (Chapters 1, 2 & 3). First of all, SDMs have proven to better assess the distribution of species when including infra-specific information (Chapter 1) converting them on particularly powerful tools for conservation management of genetic resources (Chapter 2). In addition, we have provided an approach to integrate both, abiotic and biotic factors, in order to more accurately predict exposure to climate change (Chapter 3). Lastly, we have employed a solid method to account for the uncertainty derived from algorithm selection (Chapter 2) and the wide range of future climate predictions (Chapters 2 & 3). On the other hand, we developed a novel approach to assess environmental (based on SDMs) and geographical (based on geometric indexes) marginality of populations as these may play an essential role in the adaptation processes of species to climate change.

However, the integration of genetic information into SDMs is only in its starting point and more ambitious goals than separately modelling genetic groups could be obtained. In fact, we should evolve from the traditional use of SDMs - to indicate environmental suitability for a species or genotype – to approaches modelling fitness related traits or the frequency of alleles - that have been previously detected to be environmentally driven - in a spatially explicit framework. In these cases, the dependant variable could change from binomial (presence/absence) to continues (ranging between 0-1) representing, for instance, survival or the probability of finding a specific allele. These kind of approaches would enable us to know the distribution of relevant traits or/and genes in space. Although information availability is the major constrain for these studies, common garden experiments constitute a very valuable alternative, as they test different populations in contrasted environments. In addition, the development of next generation sequencing techniques is resulting in genetic datasets increasingly becoming available.

Conclusions

Conclusions

1. Species distribution models, commonly based on abiotic factors, improve their predictions when incorporating infra-specific level information, *i.e.* genetic data, providing useful tools to evaluate species' vulnerability to climate change and thereby to assess conservation forest management (Chapters 1 & 2).
2. When assessing vulnerability to climate change, it is possible to integrate abiotic and biotic exposure in order to guide management for adaptation (Chapter 3).
3. Assessing exposure to climate change requires dealing with future climate uncertainties derived from the wide range of future climate predictions available. To circumvent these uncertainties, approaches integrating all the available information, should be favoured in regards to those dealing with a single or averaged climate prediction (Chapter 2 & 3).
4. Evaluating abiotic and biotic exposures to climate change can guide the selection of appropriate management and conservation strategies (Chapters 2 and 3). Based on exposure assessment, diverse options are recommended to enhance adaptation, differentiating among *in situ* conservation, *ex situ* conservation or *in situ* conservation with monitoring (Chapter 2), and between breeding programmes or silvicultural methods (Chapter 3).
5. Exposure evaluation should be included as a requirement for dynamic conservation units (DCUs) within EUFGIS, the most important European conservation network of forest genetic resources. DCUs should adequately represent all genetically defined clades of the species included in the network. (Capítulo 2)
6. Marginal populations may play an important role in species' adaptation. To understand how their genetic patterns and subsequent fitness are affected, it is necessary to define marginality in a standardized procedure based on major processes affecting gene flow, *i.e.* isolation by distance (geography) and isolation by environment (environment). The former can be approached by geometric criteria while SDMs can provide a tool to assess the latter (Chapter 4).

7. Integrating geographic and environmental aspects of marginality may enable to disentangle their separate and ensemble effects on genetic diversity. It would be interesting to compare core *vs* marginal populations genetic patterns to obtain solid conclusions, although this is not always possible as population sampling has been majorly directed by geographic factors (Chapter 4).
8. The genetic structure within *Pinus pinaster* Ait. is primarily a consequence of its demographic history. Nevertheless, local environmental adaptation played a role in shaping its lower order phylogeographical structure suggesting that niche evolution contributed to the differentiation of the clades (Chapter 1).
9. *Pinus pinaster* Ait will likely be largely affected by climate change (Chapters 2 and 3), particularly along the Spanish Mediterranean coast (Chapter 3). *Pinus halepensis* Mill., although also affected, seems to be better able to cope with predicted climate alterations (Chapter 2)
10. According to the future climate predictions available, the suitable area of *Fusarium circinatum*, a potential fungus pathogen for *Pinus pinaster*, will be reduced in the future, thus reducing the biotic exposure of the pine, particularly in the north-western edge of the Iberian Peninsula. Still, strict monitoring should be maintained given the large negative ecological and economic impacts of an outbreak (Chapter 3).

Conclusiones

Conclusiones

1. Los modelos de distribución de especies, comúnmente basados en factores abióticos, mejoran sus predicciones al incorporar información infra-específica, es decir, datos genéticos, y constituyen herramientas útiles para evaluar la vulnerabilidad de las especies al cambio climático y por tanto para guiar la gestión forestal (Capítulos 1 & 2).
2. De cara a evaluar la exposición al cambio climático, es necesario considerar, además de la exposición abiótica, la biótica, para guiar la gestión forestal para la adaptación (Capítulo 3).
3. Evaluar la exposición al cambio climático requiere lidiar con la incertidumbre derivada del gran rango de predicciones de clima futuro disponibles. Para evitar esta incertidumbre, se deben favorecer enfoques que integren toda la información disponible, evitando el empleo de una única predicción o de una predicción media (Capítulos 2 & 3).
4. La selección de estrategias de manejo y conservación apropiadas puede guiarse a través de una evaluación de la exposición al cambio climático tanto abiótica como biótica (Capítulos 2 y 3). Basándose en esta evaluación, se proponen diferentes alternativas diferenciando entre conservación *in situ*, conservación *ex situ* y conservación *in situ* con monitorización (Capítulo 2) o entre utilizar programas de mejora genética o métodos silvícolas (Capítulo 3).
5. La evaluación de la exposición al cambio climático se debe incorporar como un requisito en las unidades de conservación dinámicas (UCDs) de EUFGIS, la red Europea más importante para la conservación de recursos genéticos forestales. Las UCDs deben representar adecuadamente todos los grupos genéticos definidos para las especies incluidas en la red (Capítulo 2).
6. Las poblaciones marginales pueden jugar un papel importante en la adaptación de las especies. Para entender la relación entre sus patrones genéticos y su consiguiente fitness se ven afectados, es necesario definir la marginalidad de manera estandarizada y basada en los principales procesos que afectan al flujo genético, es decir, aislamiento por distancia y aislamiento por ambiente. El primer proceso se puede evaluar a partir de criterios

genéticos, mientras que los MDEs pueden emplearse para evaluar el segundo (Capítulo 4)

7. La integración de los aspectos geográficos y ambientales de la marginalidad puede permitir desentrañar sus efectos individuales y conjuntos sobre la diversidad genética. Sería interesante comparar los patrones genéticos de poblaciones centrales *vs* marginales aunque esto no siempre es posible ya que el muestreo suele estar dirigido principalmente por criterios geográficos (Capítulo 4).
8. La estructura genética de *Pinus pinaster* Ait. es principalmente consecuencia de su historia demográfica. Sin embargo, la adaptación local ha jugado también un papel importante en la determinación de su estructura filogeogeográfica sugiriendo que la evolución de nicho ha contribuido a la diferenciación de los grupos genéticos (Capítulo 1).
9. El cambio climático parece tener importantes consecuencias sobre *Pinus pinaster* Ait (Capítulos 2 & 3), particularmente a lo largo de la costa Mediterránea española. Aunque *Pinus halepensis* Mill., también se verá afectado por las alteraciones climáticas, parece estar mejor preparado para lidiar con ellas (Capítulo 2).
10. Según las predicciones climáticas futuras, el área climáticamente apta para el hongo *Fusarium circinatum*, potencialmente patógeno para *Pinus pinaster* Ait.*,* se reducirá, reduciéndose por tanto la exposición biótica del pino, particularmente en el noroeste de la Península Ibérica. Es necesario mantener monitorizadas las poblaciones en estas zonas dadas las graves consecuencias ecológicas y económicas de una epidemia.

Annex

- Supporting information Chapter 1
- Supporting information Chapter 2
- Supporting information Chapter 3
- Supporting information Chapter 4

Supporting information Chapter 1:

APPENDIX S1:

Detailed description of genetic clades

APPENDIX S2:

Specifications for the preparation of presence/pseudoabsence data inputs, algorithm settings for the species distributions models, comparisons of model performance and scatterplots of phylogenetic-environmental distances /phylogenetic-geographic distances

APPENDIX S3:

Detailed description of bioclimatic variables and selection procedure

APPENDIX S4:

Geographic projections of the five algorithms (GLM, GAM, RF, CTA, MaxEnt and Ensemble) of the species and its genetic clades (G1-G8).

APPENDIX S1:

Detailed description of genetic clades

Pinus pinaster phylogeography has been widely studied using a variety of molecular markers, each of them unveiling part of the species' evolutionary history. Maternally inherited markers (mitochondrial DNA sequences), which reflect seed dispersal events, pointed to three main clusters, namely a Moroccan cluster, a western cluster (Iberian Peninsula and continental France, except Catalonia and southeastern France), and an eastern cluster (southeastern France, Catalonia, Corsica, Italy, Sardinia, Tunisia and Algeria) (Burban & Petit, 2003). Paternally inherited markers (chloroplast Simple Sequence Repeats (SSRs)), which reflect pollen dispersal events, identified eight main genetic clusters (Vendramin *et al.*, 1998; Bucci *et al.*, 2007), namely northern Spain and France, forming an Atlantic subgroup; the Atlantic, central, and eastern regions of the Iberian Peninsula, forming a central subgroup; and western and eastern Africa as well as northern Italy, Corsica, and Sardinia, forming a southeastern subgroup.

A first set of bi-parental markers consisting of 12 nuclear SSRs (Santos-del-Blanco *et al.*, 2012) identified five genetically-defined regions: Morocco, Corsica, the Atlantic coast of France, as well as the Atlantic and Mediterranean regions of the Iberian Peninsula. The delimitation of the same genetic pools was confirmed by different sets of bi-parental markers (nine nuclear SSRs and 1745 Single Nucleotide Polymorphisms - SNPs) further dividing the Mediterranean Spain region into central Spain and south-eastern Spain (Jaramillo-Correa *et al.*, 2015). Here, by means of 12 nSSRs (analysed in Jaramillo-Correa *et al.*, 2015) and 266 SNPs (subset of the 384 SNPs analysed in Jaramillo-Correa *et al.*, 2015) we further confirmed the delimitation of the same six regions with an even higher resolution, especially in the Iberian Peninsula, where three gene pools can be further defined based on the estimated membership coefficients for each individual within each cluster (*Q;* see Table S1). *Q* were calculated as a mean-over-population of individual coefficients of membership to each of the eight clusters defined by Structure clustering method (Pritchard *et al.*, 2000) on the 266 SNP data set. The cluster harbouring the highest *Q* value for each population (ranging from 0.38 to 0.99) was one of the parameters used to define genetic clusters, along with highly concordant Unweighted Pair Group Method with Arithmetic Mean (UPGMA) phylogenetic results. The southern Spain group (populations COM, CAZ, ORI, QUA) did not appear here as a separate cluster, but as a mixed group sharing characteristics of Moroccan and eastern Spanish groups.

The combination of information from all available molecular markers permitted defining eight genetic clades that characterize the evolutionary history of *P. pinaster*, described in detail below:

G1 – an Atlantic Iberian Peninsula clade defined by northwest Spain and Portugal. Both chloroplast and nuclear markers detected a substructure within the Atlantic populations that subdivide in Atlantic France (G3) and Atlantic Iberian Peninsula (G1).

G2 – an eastern clade including southeastern France, northern Italy and Corsica. All markers point to this clade as an eastern origin, even though only nuclear markers for Corsica populations are available. Populations from this cluster are independent from all other populations.

G3 – an Atlantic France clade formed by all Atlantic French populations. Both chloroplast and nuclear markers detect a substructure within the Atlantic populations that subdivide them into Atlantic France (G3) and Atlantic Iberian Peninsula (G1).

G4 – a Moroccan clade formed by all Moroccan populations. Although nuclear markers point to a certain degree of connection between Moroccan and southern Spanish populations, mitochondrial and chloroplast markers point to the Strait of Gibraltar as a genetic breakpoint (with the exception of the northern Moroccan population, Punta Cires; see Burban & Petit, 2003; Bucci *et al.*, 2007).

G5, G6, G7 – represent, respectively, eastern, central and southern Spanish clades. This substructure of the Spanish populations was partially detected with chloroplast markers, and was better defined with nuclear markers. The eastern clade G5 results from a genetic admixture between the southern and central Spanish clade. The central clade G6 results from an admixture between the Atlantic clades and the other Spanish clades. Finally, the southern clade, G7 results from an admixture between the Moroccan clade and the Spanish clades (see Table S1).

G8 – a Tunisian clade. This clade stands by itself, based on mitochondrial and chloroplast markers that point to its eastern origin, while nuclear markers connect it to an Atlantic clade. However, the origin of the Tunisian populations is dubious in the dataset screened with nuclear data. We suspect that these individuals were mistakenly labelled as of Tunisian origin for nSSR and SNPs studies, as both analysed the same set of individuals (Jaramillo-Correa *et al.*, 2015). As Tunisian individuals, we would have expected them to group with eastern European clades based on mitochondrial and chloroplast markers for which distinct sets of individuals were analyzed (see Vendramin *et al.*, 1998; Burban & Petit, 2003; Bucci *et al.*, 2007).

Table S1: Estimated membership coefficients (Q) for each individual within each cluster. In bold, highest Q values used for assignment of each population to a cluster. Colours correspond to the overall genetic clusters eventually defined for the study.

Population	**Coordinates**	***Q* for 8 Structure clusters**							
		Cl. 1	**Cl. 2**	**Cl. 3**	**Cl. 4**	**Cl. 5**	**Cl. 6**	**Cl. 7**	**Cl. 8**
TBK		0.01	0.02	0.14	0.03	**0.60**	0.05	0.04	0.11
SID		**0.98**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TAM		**0.99**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MAD		**0.87**	0.00	0.01	0.00	0.00	0.08	0.00	0.03
FCN		0.01	0.01	0.03	0.01	0.02	0.03	**0.87**	0.02
COM		**0.27**	0.01	0.04	0.02	0.03	**0.48**	0.01	**0.13**
CAZ		**0.12**	0.01	0.05	0.02	0.03	**0.46**	0.08	**0.25**
ORI		**0.27**	0.01	0.01	0.01	0.01	**0.47**	0.02	**0.20**
QUA		**0.06**	0.02	0.03	0.01	0.04	**0.61**	0.03	**0.21**
ESL		0.01	0.01	0.03	0.02	0.01	**0.62**	0.02	**0.29**
CAL		0.01	0.01	0.08	0.05	0.02	**0.54**	0.04	**0.25**
OLB		0.04	0.02	0.30	0.05	0.02	**0.39**	0.02	**0.15**
BON		0.09	0.01	0.10	0.10	0.00	**0.50**	0.05	**0.15**
SIN		0.03	0.01	0.21	0.03	0.02	**0.43**	0.04	**0.22**
BAY		0.01	0.01	**0.66**	0.07	0.03	0.07	0.01	0.14
SAL		0.01	0.02	**0.68**	0.05	0.05	0.08	0.04	0.08
ARN		0.00	0.02	**0.38**	0.03	0.09	0.19	0.05	0.24
CEN		0.00	0.01	**0.55**	0.01	0.07	0.12	0.04	0.20
VAL		0.00	0.00	**0.69**	0.05	0.02	0.08	0.02	0.13
COC		0.02	0.07	**0.54**	0.11	0.05	0.08	0.02	0.11
CAR		0.02	0.02	**0.60**	0.07	0.07	0.05	0.03	0.14
CUE		0.01	0.00	**0.57**	0.17	0.07	0.04	0.02	0.11
TBY		0.01	0.01	**0.58**	0.01	0.03	0.09	0.02	0.26
LEI		0.01	0.02	0.13	0.02	**0.66**	0.04	0.02	0.10
SAC		0.01	0.01	0.14	0.02	**0.74**	0.03	0.02	0.03
SEG		0.00	0.01	0.06	0.03	**0.86**	0.02	0.01	0.02
ARM		0.00	0.00	0.01	0.01	**0.96**	0.01	0.00	0.01
ALT		0.00	0.01	0.01	0.04	**0.83**	0.02	0.00	0.08
ROD		0.00	0.01	**0.54**	0.02	0.03	0.13	0.02	0.26
LAM		0.00	0.01	0.01	0.01	**0.95**	0.00	0.01	0.00
CAD		0.00	0.00	0.01	0.03	**0.93**	0.01	0.01	0.01
SIE		0.01	0.02	0.07	0.06	**0.66**	0.08	0.01	0.09
PUE		0.00	0.01	0.01	0.01	**0.93**	0.01	0.01	0.02
CAS		0.00	0.01	0.06	0.01	**0.79**	0.06	0.03	0.04
BIB		0.00	**0.71**	0.01	0.07	0.04	0.04	0.01	0.13
PIA		0.00	**0.81**	0.02	0.01	0.08	0.02	0.02	0.04
PIE		0.00	**0.84**	0.04	0.00	0.01	0.03	0.01	0.06
PIN		0.00	**0.81**	0.01	0.07	0.01	0.02	0.01	0.07
PLE		0.00	0.00	0.05	**0.85**	0.02	0.02	0.02	0.03
STJ		0.01	0.01	0.02	**0.84**	0.03	0.02	0.02	0.03
OLO		0.01	0.01	0.03	**0.84**	0.03	0.02	0.04	0.02
VER		0.01	0.01	0.05	**0.80**	0.05	0.02	0.02	0.04
HOU		0.01	0.01	0.03	**0.91**	0.01	0.02	0.01	0.02
MIM		0.00	0.01	0.06	**0.80**	0.02	0.03	0.02	0.06
PET		0.01	0.02	0.04	**0.79**	0.04	0.03	0.03	0.03

References

Bucci, G., González-Martínez, S.C., Le Provost, G., Plomion, C., Ribeiro, M.M., Sebastiani, F., Alía, R. & Vendramin, G.G. (2007) Range-wide phylogeography and gene zones in Pinus pinaster Ait. revealed by chloroplast microsatellite markers. *Molecular Ecology*, **16**, 2137–2153.

Burban, C. & Petit, R.J. (2003) Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. *Molecular Ecology*, **12**, 1487–1495.

Jaramillo-Correa, J., Rodríguez-Quilón, I., Grivet, D., Lepoittevin, C., Sebastiani, F., Heuertz, M., Garnier-Géré, P.-P., Alía, R., Plomion, C., Vendramin, G.G. & González-Martínez, S.C. Molecular proxies for climate maladaption in a long-lived tree (*Pinus pinaster* Aiton, Pinaceae). *Genetics,* **199,** 1-15.

Pritchard, J.K., Stephens, M., Donnelly, P. (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, **155**, 945-959.

Santos-del-Blanco, L., Climent, J., González-Martínez, S. & Pannell, J.R. (2012) Genetic differentiation for size at first reproduction through male versus female functions in the widespread Mediterranean tree Pinus pinaster. *Annals of Botany*, **110**, 1449–1460.

Vendramin, G.G., Anzidei, M., Madaghiele, A. & Bucci, G. (1998) Distribution of genetic diversity in Pinus pinaster Ait. as revealed by chloroplast microsatellites. *Theoretical and Applied Genetics*, **97**, 456–463.

APPENDIX S2:

Specifications for the preparation of presence/pseudo-absence data inputs, algorithm settings for the species distribution models, comparison of model performance and scatterplots of phylogenetic-environmental distances / phylogenetic-geographic distances

Preparation of presence/pseudo-absence data inputs

The Tree Species Distribution for Europe (TSDE; Köble & Seufert, 2001) from the Joint Research Centre's (JRC) AFOLU data portal[1] and (ii) the EUFORGEN database from the European forest genetic resources programme[2] were combined to approximately obtain the natural distribution of *Pinus pinaster*.

In general, the distribution of maritime pine is similar in both databases (Fig. S1). There are, however, some areas of mismatch in which TSDE detects presences not included within EUFORGEN limits. These areas are located mainly in southwestern Spain, northwestern France and the Croatian coast, and are probably due to plantations (*i.e.* non-natural species presence). The opposite disagreement also occurs, with EUFORGEN indicating presence while TSDE does not (mainly in the Italian coast and southern Spain). EUFORGEN has been created at a broader scale than TSDE, while the latter provides more information about the exact location of the species' presence within the polygon defined by EUFORGEN.

TSDE was reclassified considering 0 as a pseudo-absence cell, and any percentage higher than 0 as a presence cell. In the European territory, presence records were selected as those cells that are considered as presences by TSDE and that are also included within the EUFORGEN shape limits. On the other hand, as TSDE does not extend to Northern Africa, only EUFORGEN was therefore used with all cells within the limits of this database, as assumed to represent true presences.

---

[1] http://afoludata.jrc.ec.europa.eu/index.php/public_area/tree_species_distribution

[2] http://www.euforgen.org/

Presences of genetic clades were defined as the subset of the overall presences records that belong to one specific genetic clade. Possible pseudo-absences correspond to all the rest of the territory including presences from other genetic clades. One of the populations included in the SNP analysis (Fuencaliente, FCN, Spain) was not included within EUFORGEN limits and was not considered in further analyses.

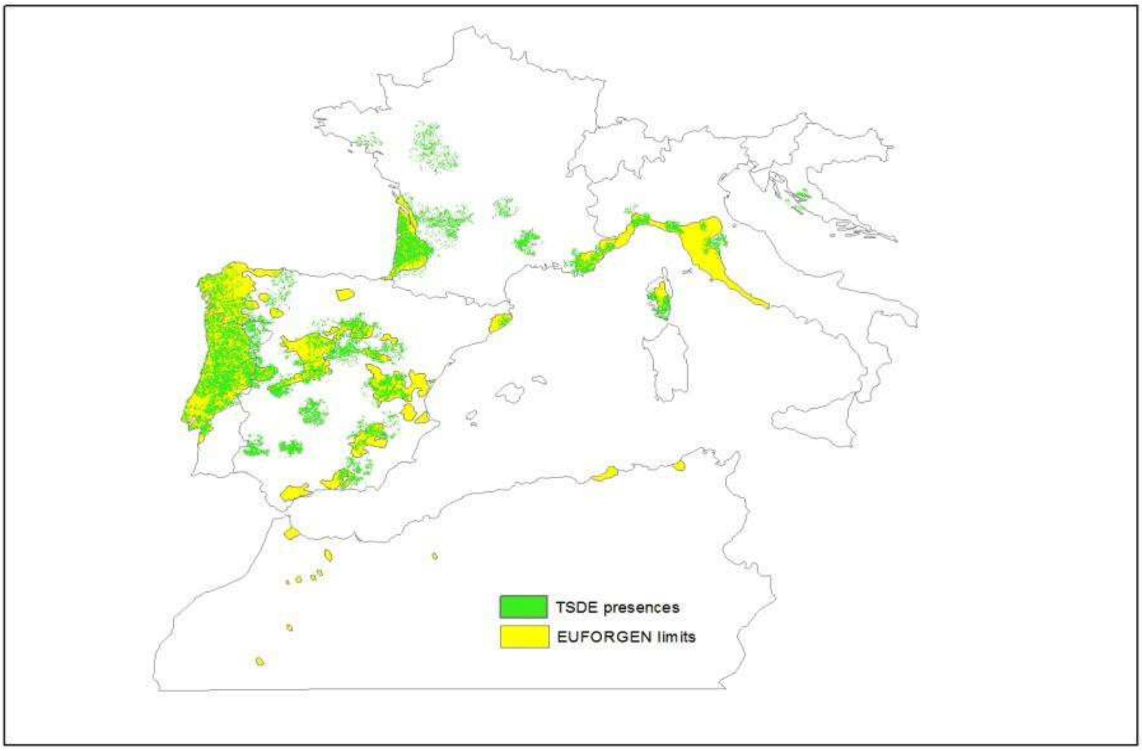


Figure S1: Combination of EUFORGEN and "Tree Species Distribution for Europe" (TSDE) data bases for *Pinus pinaster* distribution. TSDE has been reclassified considering cells with percentage higher than 0 to represent presence of the species.

The final number of presences selected for the species model summed 128,653, while the number of presences for the eight genetic clades amounted to: 59,109 (G1), 12,215 (G2), 17,877 (G3), 4,940 (G4), 8,785 (G5), 13,005 (G6), 4,981 (G7), and 3,997 (G8). The patches corresponding to unclassified genetic clades were not included in the individual calibrations of genetic clades, but were considered when modelling the species-level model. Number of pseudo-absences is specified in the next section, as it depends on algorithm selection.

Algorithm specifications

Five different statistical methods were individually calibrated and then combined to build an ensemble Species Distribution Model (SDM). As regression methods, we selected Generalized Linear Models (GLM; McCullagh & Nelder, 1989) and Generalized Additive Models (GAM; Hastie & Tibshirani, 1990). Random Forest (RF; Breiman, 2001) represents machine-learning methods while Classification Tree Analysis (CTA; Breiman, 1984) stands for classification methods. Finally, we also used MaxEnt representing a maximum entropy model (Phillips *et al.*, 2006). We set up GLM to perform a stepwise variable reduction/addition procedure by maximizing Akaike Information Criterion (AIC) starting from a full model based on linear and quadratic terms for all five variables. The number of trees was fixed to 1,000 for RF while default settings were applied in GAM, CTA and MaxEnt. Model performance was assessed from the True Skill Statistic (TSS; Allouche *et al.*, 2006), the Area Under the ROC Curve (AUC) (Fielding & Bell, 1997) and the H-Measure (Hand, 2010, 2012). Values of H-Measure are calculated from binary models constructed from probabilistic outputs by defining thresholds that optimized TSS values.

Concerning pseudo-absence selection, given the broad spatial and environmental scale of *P. pinaster*'s range, a large number of pseudo-absences were needed. Randomly selecting pseudo-absences produce the most accurate GLM and GAM, while RF and CTA show less variation according to pseudo-absence selection methods (Barbet-Massin *et al.,* 2012). Large number of pseudo-absences records is also recommended for GLM and GAM (Barbet-Massin *et al.,* 2012), and this is consistent with the results obtained for MaxEnt (Phillips & Dudík, 2008). We set the number of randomly selected pseudo-absences to five times the number of presences detected in the species model. For the clade models, we maintained the number of pseudo-absences used in the species model. Pseudo-absences and presences were given weights inversely proportional to their numbers, so as to give equal total weights to the two sets, as recommended by Barbet-Massin *et al.* (2012).

The large number of occurrence records available permitted a random division of each dataset (corresponding to the species and to each clade) into two equally sized subsets for training and evaluating maintaining the initial proportion between presence and pseudo-absence records. Subsequently, during the model evaluation stage, we only used 20 % of the set-aside evaluation records to minimize computation time. The reduced evaluation subset contained equal numbers of presence and pseudo-absence records.

Comparison of model performance

We built 45 different models, five for each of the eight genetic clades, and another five for the species model.

The models calibrated individually for each genetic clade reached better performance (slightly superior AUC and TSS scores and notably higher H-Measure values) than the models calibrated at the species level (Table S2). Although the five statistical algorithms performed well (with AUC, TSS and H-Measure scores all higher than 0.968, 0.851 and 0.767, respectively, in the individually calibrated clade models), RF displayed the best AUC, TSS and H-Measure scores compared to the other algorithms in all cases except for G8, where GLM and GAM performed slightly better.

The ensemble models summarize the individual statistical model information for each clade and for the species into one single prediction. Overall, TSS and H-Measure scores of ensemble models were higher, compared to most models from individual statistical algorithms (Table S3). Nevertheless, TSS and H-Measure scores from RF generally outcompete ensemble TSS and H-Measure scores (see Table S3 vs. Table S2). Sensitivity and specificity scores obtained in the final ensemble binary models are also considerably high (above 89 % in all cases), and again, genetic clade's ensemble models outperformed the species ensemble model.

Table S2: Assessment of model performance by means of True Skill Statistic (TSS), Area Under the Curve (AUC) and H-Measure scores. Values are displayed separately according to model inputs (species level models and genetic clade level (G1-G8) models) and, additionally, to the five different algorithms employed: General Linear Model (GLM), General Additive Model (GAM), Random Forest (RF), Classification Tree Analysis (CTA) and MaxEnt. Values of H-Measure are calculated from binary models constructed from probabilistic outputs by defining thresholds that optimized TSS values. This evaluation is based on a split sample test.

Model	Algorithm	AUC score	TSS score	H-Measure
Species Level	GLM	0.893	0.634	0.455
	GAM	0.909	0.666	0.495
	RF	0.980	0.877	0.804
	CTA	0.934	0.773	0.647
	MaxEnt	0.909	0.643	0.465
G1	GLM	0.995	0.950	0.919
	GAM	0.997	0.967	0.947
	RF	1.000	0.992	0.988
	CTA	0.992	0.969	0.950
	MaxEnt	0.996	0.961	0.938
G2	GLM	0.979	0.919	0.871
	GAM	0.983	0.942	0.908
	RF	0.999	0.955	0.928
	CTA	0.980	0.951	0.922
	MaxEnt	0.987	0.920	0.873
G3	GLM	0.999	0.988	0.981
	GAM	1.000	0.996	0.993
	RF	1.000	0.996	0.994
	CTA	0.994	0.986	0.978
	MaxEnt	0.999	0.988	0.981

G4	GLM	0.968	0.851	0.767
	GAM	0.981	0.933	0.893
	RF	0.999	0.953	0.925
	CTA	0.979	0.948	0.917
	MaxEnt	0.986	0.904	0.846
G5	GLM	0.996	0.980	0.968
	GAM	0.997	0.983	0.973
	RF	1.000	0.988	0.981
	CTA	0.991	0.982	0.971
	MaxEnt	0.997	0.976	0.962
G6	GLM	0.990	0.949	0.919
	GAM	0.993	0.960	0.936
	RF	0.999	0.977	0.963
	CTA	0.986	0.964	0.943
	MaxEnt	0.994	0.943	0.909
G7	GLM	0.989	0.929	0.887
	GAM	0.989	0.954	0.926
	RF	1.000	0.966	0.946
	CTA	0.983	0.962	0.939
	MaxEnt	0.992	0.939	0.903
G8	GLM	1.000	0.996	0.993
	GAM	1.000	0.996	0.993
	RF	1.000	0.987	0.980
	CTA	0.994	0.990	0.984
	MaxEnt	0.999	0.991	0.985

Table S3: True Skill Statistic (TSS), H-Measure, Cutoff, Sensitivity and Specificity scores of the Ensemble model (average of the five algorithms), for the species level model and for the different genetic clades' (G1-G8) models. This evaluation is based on a split sample test.

Model	**H-Measure**	**TSS**	**Cutoff value**	**Sensitivity**	**Specificity**
Species	0.692	0.804	456	90.914	89.450
G1	0.967	0.979	488	99.132	98.799
G2	0.923	0.952	458	99.099	96.070
G3	0.991	0.994	491	99.720	99.720
G4	0.936	0.960	494	98.583	97.436
G5	0.977	0.986	455	99.848	98.748
G6	0.953	0.971	494	99.052	97.975
G7	0.929	0.956	569	97.410	98.160
G8	0.985	0.991	607	99.083	100.000

Scatterplots of phylogenetic-genetic distances / phylogenetic-geographic distances

Figure S2 shows the scatterplots of phylogenetic *vs* genetic distances and phylogenetic *vs* geographic distances. Each dot represents a pairwise distance among the eight genetic clades. The scatterplots suggest a positive correlation between genetic and environmental distances and, although less obvious, the same pattern can be observed for genetic and geographic distances. This is corroborated by the Mantel regression, which explained 12.6 % of the variation in phylogenetic distance among the eight clades by means of linear combinations of environmental and geographic distance. In this regression, the linear term for environmental distance was barely significant (p = 0.06), while the intercept (p = 0.87) and the geographic distance (p = 0.74) were not.

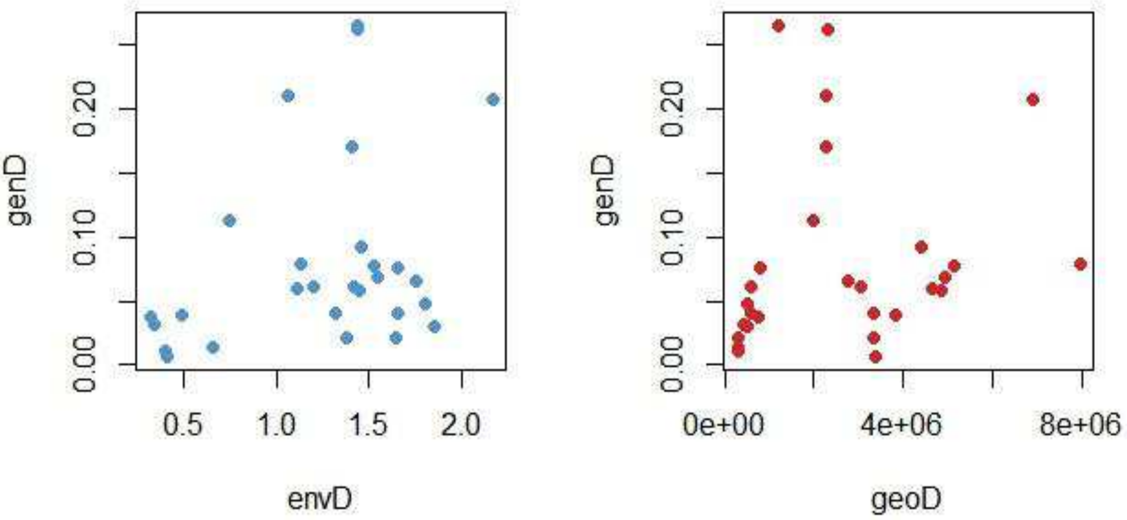


Figure S2: Scatterplots of phylogenetic *vs* environmental distances (left side) and phylogenetic *vs* geographic distances (right side).

References

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.

Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **3**, 327–338.

Breiman, L. (2001) Random forests. *Machine Learning*, **45**, 5–32.

Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. (1984) *Classification and Regression Trees*, Wadsworth & Brooks/Cole Advanced Books & Software, Monterey.

Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.

Hand, D.J. (2012) Assessing the performance of classification methods. *International Statistical Review*, **80**, 400–414.

Hand, D.J. (2010) Evaluating diagnostic tests: The area under the ROC curve and the balance of errors. *Statistics in Medicine*, **29**, 1502–1510.

Hastie, T.J. & Tibshirani, R. (1990) *Generalized additive models*, Chapman & Hall/CRC, Boca Raton - London - New York - Washington D.C.

Köble, R. & Seufert, G. (2001) *Novel maps for forest tree species in Europe. Proceedings of the 8th European Symposium on the Physico-Chemical Behaviour of Air Pollutants: "A Changing Atmosphere!,"* pp. 17–20. Data set available online at: http://afoludata.jrc.ec.europa.eu/index.php/dataset/detail/66, Torino.

McCullagh, P. & Nelder, J.A. (1989) *Generalized linear models*, (ed. by C. and Hall).

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Phillips, S.J. & Dudík, M. (2008) Modeling of species distributions with Maxent : new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.

APPENDIX S3:

Detailed description of bioclimatic variables and selection procedure

We considered 23 bioclimatic variables for the analysis, representative of the period 1950-2000. Nineteen of them (BIO1 to BIO19; see Box 1) were downloaded from the current bioclimatic variables available in WORLDCLIM (Hijmans *et al.*, 2005). Four new variables were created following Zimmermann *et al.* (2007), as it characterizes water availability better: summer and spring potential evapotranspiration ($ETPT_{summer}$; $ETPT_{spring}$) and summer and spring moisture index ($MIND_{summer}$; $MIND_{spring}$) (see Box 2).

In order to avoid multicollinearity effects, we retained variables with Pearson correlations lower than 0.75, as the use simple methods based on rules of thumb have proved to be as effective as more complicated methods (Dormann *et al.*, 2013). This threshold value is a bit less restrictive than the common value of 0.70, but more flexible thresholds have also been used in the literature (see Elith *et al.*, 2006 for an example). Among the highly correlated variables we kept the one with highest explained deviance scores (D^2) when individually fitted in a Generalized Linear Model (GLM; McCullagh & Nelder, 1989). We avoided the use of BIO8 and BIO9, as the steep gradient shown by these variables (*i.e.* very often two adjacent cells are characterized by extremely different values within the study area for no obvious reason) may lead to artefacts in the SDM output maps. From the remainining set we selected two temperate and two water-related variables considering that both moisture and temperature have been shown to play a key role in Mediterranean species' distribution (Benito-Garzón *et al.*, 2013). We chose BIO3 (Isothermality), BIO11 (Mean Temperature of Coldest Quarter), BIO13 (Precipitation of Wettest Month) and $MIND_{summer}$ as best predictor variables. Finally, we checked for possible collinearity problems among the selected variables by means of a Variation Inflation Factor test (VIF; Belsley, 1991; Hair Jr *et al.*, 1995), which indicated no problem since all VIF values were below a threshold value of 10.

Box 1: variables from WORLDCLIM (Hijmans *et al.*,

BIO1 = Annual Mean Temperature
BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
BIO3 = Isothermality (BIO2/BIO7) (* 100)
BIO4 = Temperature Seasonality (standard deviation *100)
BIO5 = Max Temperature of Warmest Month
BIO6 = Min Temperature of Coldest Month
BIO7 = Temperature Annual Range (BIO5-BIO6)
BIO8 = Mean Temperature of Wettest Quarter
BIO9 = Mean Temperature of Driest Quarter
BIO10 = Mean Temperature of Warmest Quarter
BIO11 = Mean Temperature of Coldest Quarter
BIO12 = Annual Precipitation
BIO13 = Precipitation of Wettest Month
BIO14 = Precipitation of Driest Month
BIO15 = Precipitation Seasonality (Coefficient of Variation)
BIO16 = Precipitation of Wettest Quarter
BIO17 = Precipitation of Driest Quarter
BIO18 = Precipitation of Warmest Quarter
BIO19 = Precipitation of Coldest Quarter

Box 2: Newly generated bioclimatic variables

Four new variables were created following Zimmermann *et al.* (2007): summer and spring potential evapotranspiration ($ETPT_{summer}$; $ETPT_{spring}$) and summer and spring moisture index ($MIND_{summer}$; $MIND_{spring}$). Potential evapotranspiration estimates water loss as a ratio depending on average temperature and solar radiation (obtained from Kumar *et al.*, 2007) following Turc's empirical equation (Turc, 1963). Moisture index is defined as the difference between precipitation (water source) and potential evapotranspiration (water loss), so values below zero indicate drought, while positive scores indicate that precipitation exceeds potential evapotranspiration.

References

Belsley, D.A. (1991) *Conditioning diagnostics: collinearity and weak data in regression*, Wiley & Sons.

Benito-Garzón, M., Ruiz-Benito, P. & Zavala, M. a. (2013) Interspecific differences in tree growth and mortality responses to environmental drivers determine potential species distributional limits in Iberian forests. *Global Ecology and Biogeography*, **22**, 1141–1151.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., Mcclean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D. & Lautenbach, S. (2013) Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 027–046.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-pereira, R., Schapire, R.E., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species ’ distributions from occurrence data. *Ecography*, **29**, 129–151.

Hair Jr, J.F., Anderson, R.E., Tatham, R.L. & William, C. (1995) *Multivariate data analysis*, Prentice Hall., New Jersey, USA.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Kumar, L., Skidmore, A.K. & Knowles, E. (2007) Modelling topographic variation in solar radiation in a GIS environment. *International Journal of Geographical Information Science*, **11**, 475–497.

McCullagh, P. & Nelder, J.A. (1989) *Generalized linear models*, (ed. by C. and Hall).

Turc, L. (1963) Evaluation des besoins en eau d’irrigation, évapotranspiration potentielle, formulation simplifié et mise à jour. *Annales agronomiques*, **12**, 13–49.

Zimmermann, N.E., Edwards, T.C., Moisen, G.G., Frescino, T.S. & Blackard, J.A. (2007) Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *Journal of Applied Ecology*, **44**, 1057–1067.

APPENDIX S4: Geographic projections of the five algorithms (GLM, GAM, RF, CTA, MaxEnt and Ensemble) of the species and its genetic clades (G1-G8).

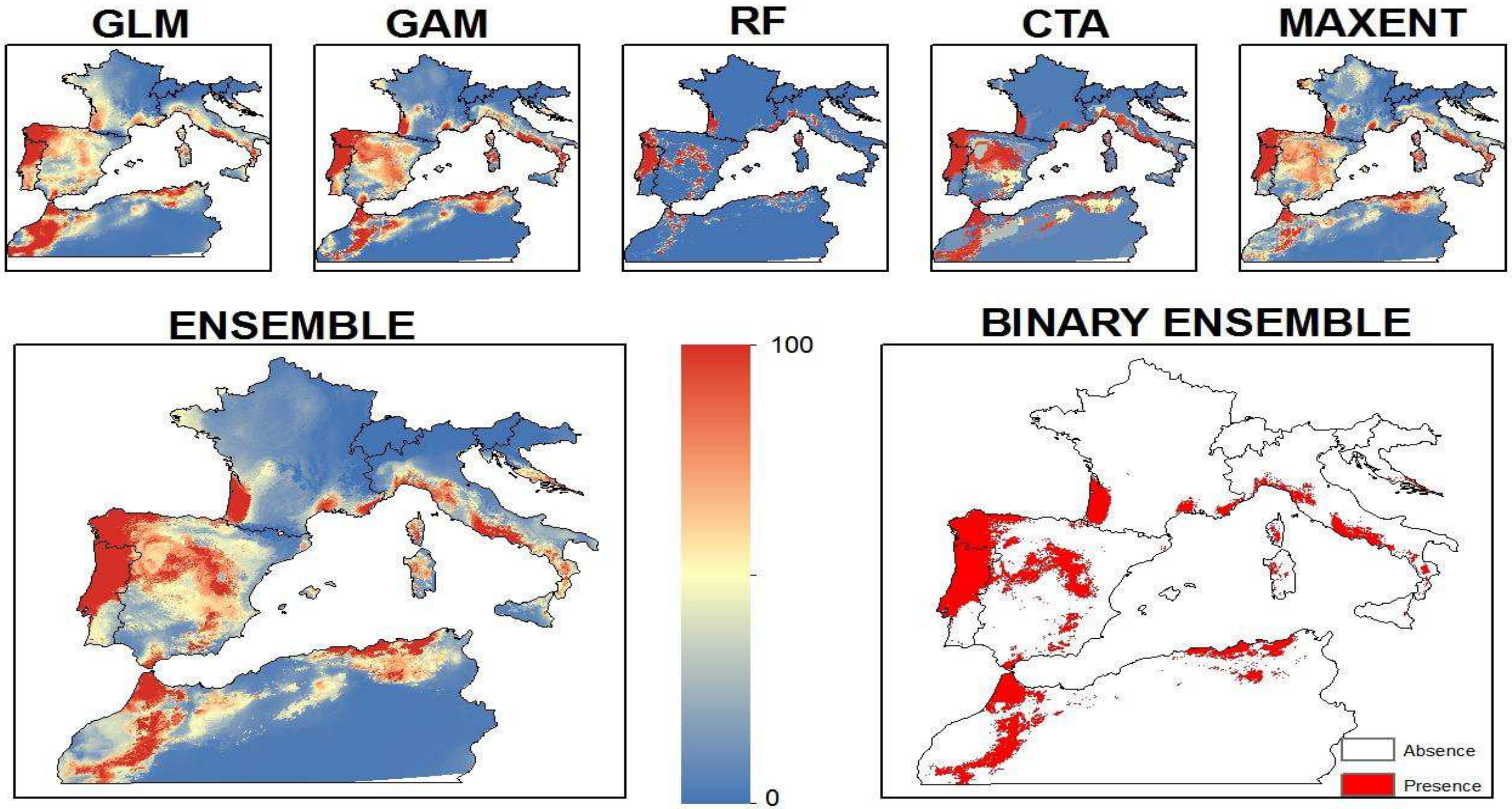


Figure S3: Geographic projection of *P. pinaster* using five algorithms (General Linear Model-GLM, General Additive Model-GAM, Random Forest-RF, Classification Tree Analysis-CTA, and MaxEnt) and the Ensemble model. Probability of presence ranges from 100 (red areas) to 0 (blue areas), while for the binary model it ranges from 1 (red areas) to 0 (white areas) (the threshold was defined optimizing TSS values from the Ensemble model, see main text for details).

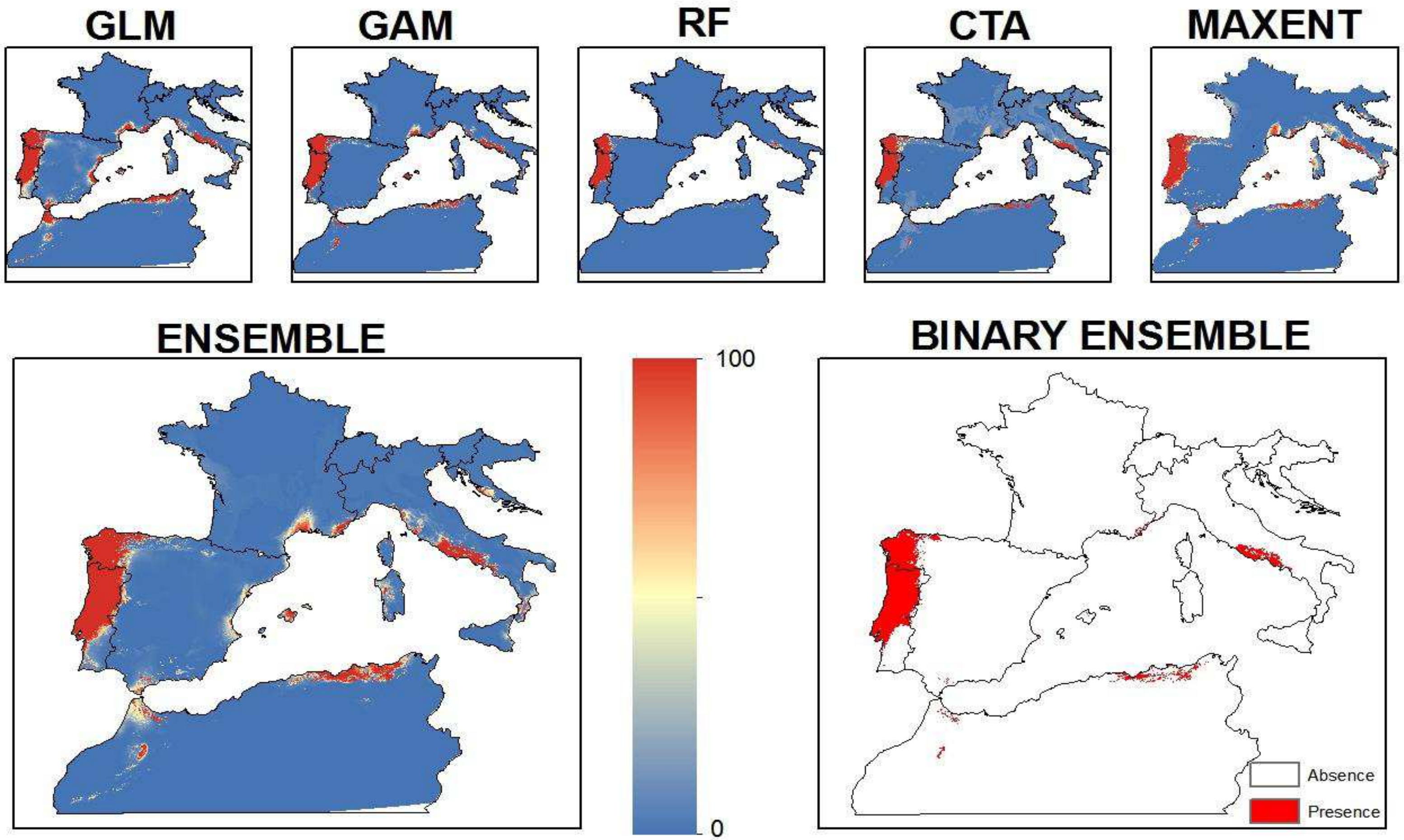


Figure S4: Geographic projection of the Atlantic Iberian Peninsula clade – G1 using five algorithms (General Linear Model-GLM, General Additive Model-GAM, Random Forest-RF, Classification Tree Analysis-CTA, and MaxEnt) and the Ensemble model. Probability of presence ranges from 100 (red areas) to 0 (blue areas), while for the binary model it ranges from 1 (red areas) to 0 (white areas) (the threshold was defined optimizing TSS values from the Ensemble model, see main text for details).

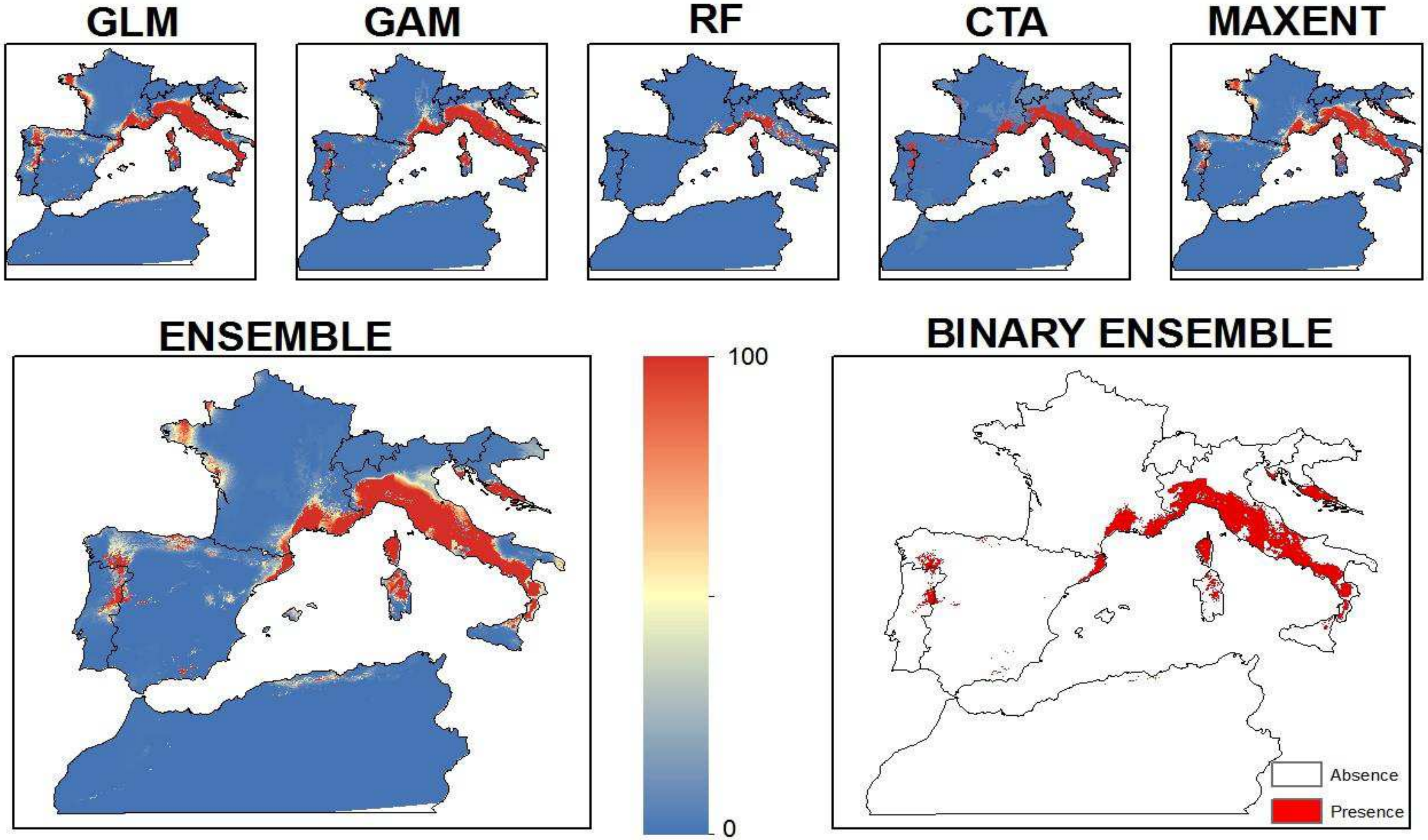


Figure S5: Geographic projection of the Eastern clade – G2 using five algorithms (General Linear Model-GLM, General Additive Model-GAM, Random Forest-RF, Classification Tree Analysis-CTA, and MaxEnt) and the Ensemble model. Probability of presence ranges from 100 (red areas) to 0 (blue areas), while for the binary model it ranges from 1 (red areas) to 0 (white areas) (the threshold was defined optimizing TSS values from the Ensemble model, see main text for details).

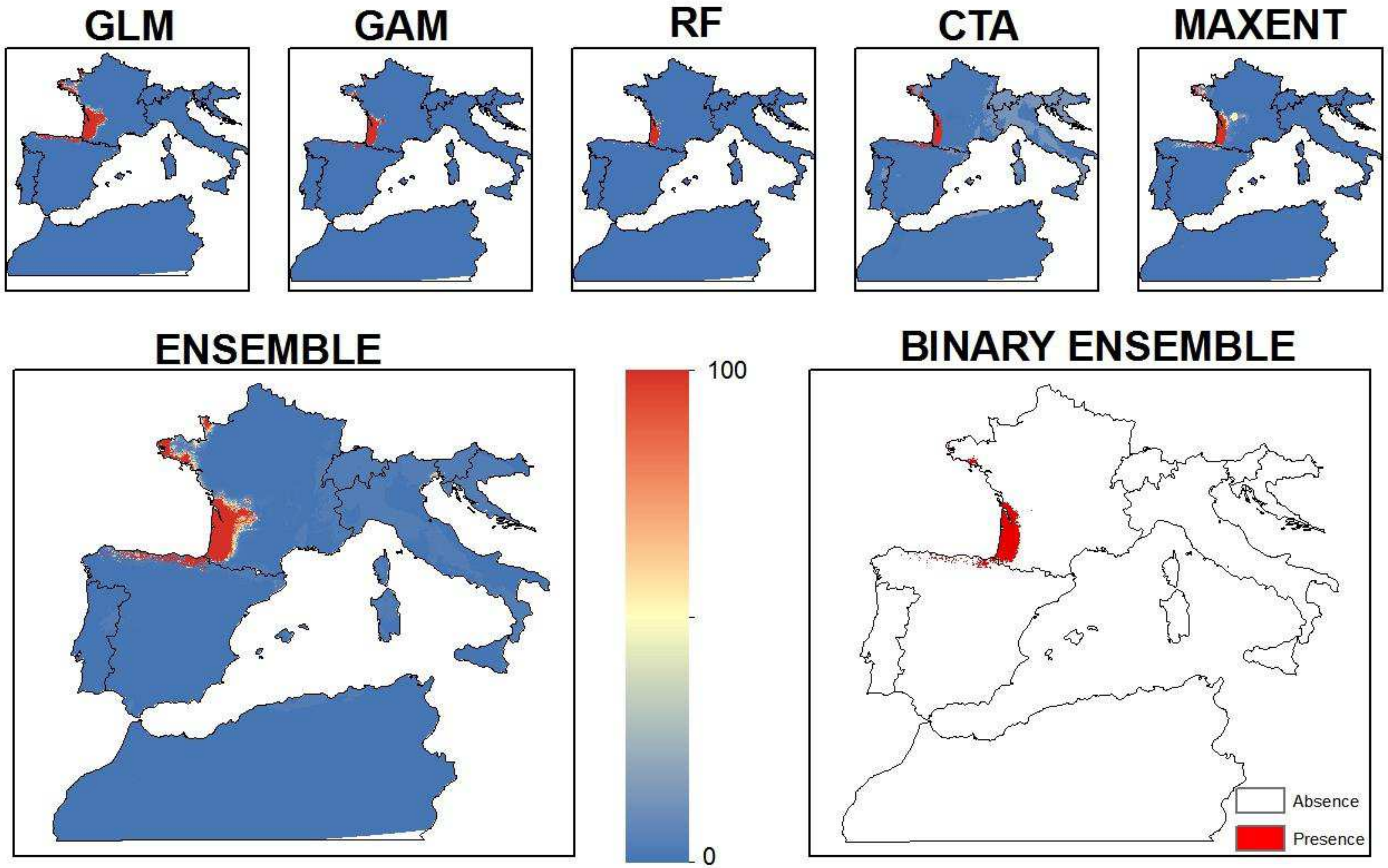


Figure S6: Geographic projection of the Atlantic France clade – G3 using five algorithms (General Linear Model-GLM, General Additive Model-GAM, Random Forest-RF, Classification Tree Analysis-CTA, and MaxEnt) and the Ensemble model. Probability of presence ranges from 100 (red areas) to 0 (blue areas), while for the binary model it ranges from 1 (red areas) to 0 (white areas) (the threshold was defined optimizing TSS values from the Ensemble model, see main text for details).

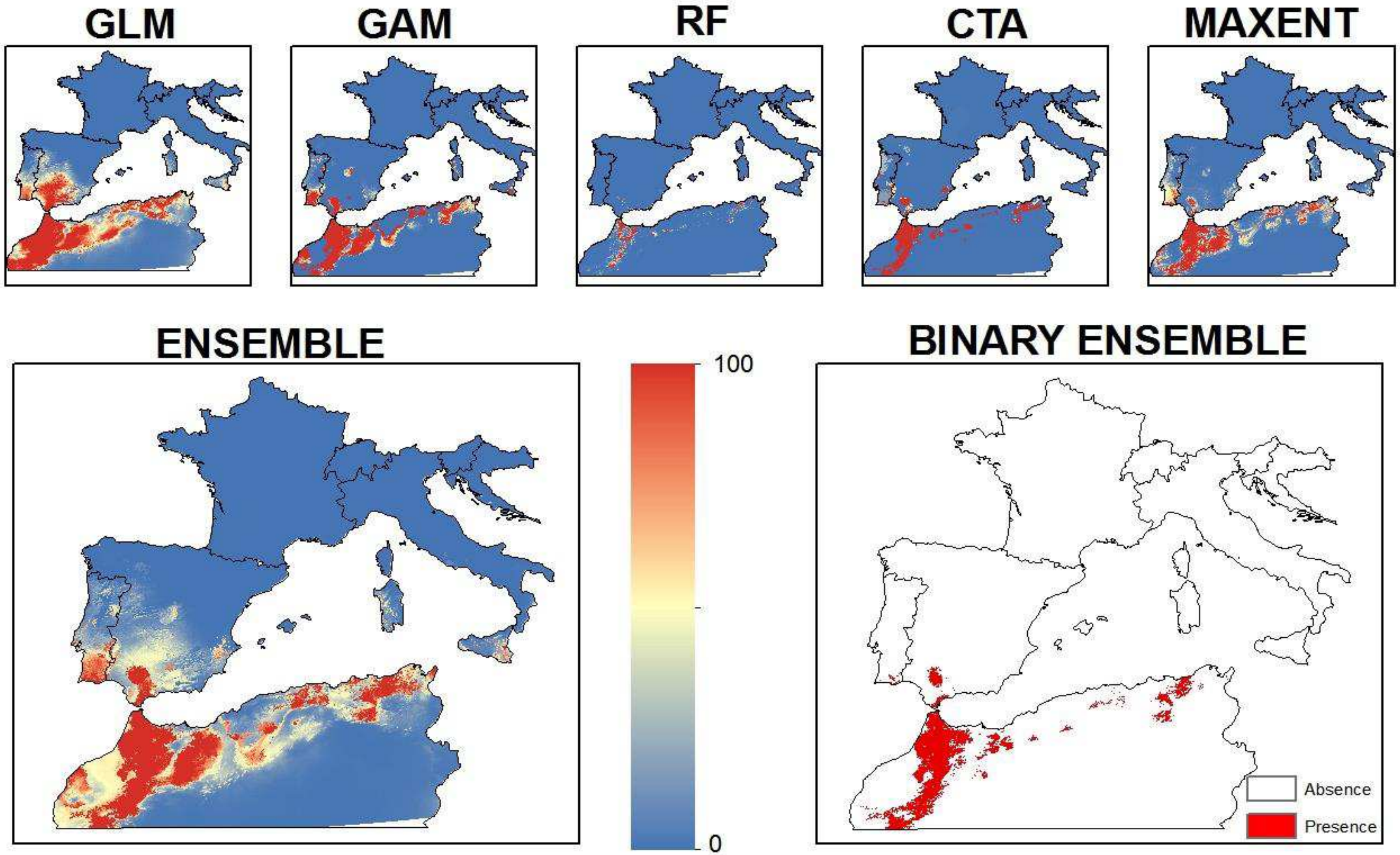


Figure S7: Geographic projection of the Moroccan clade – G4 using five algorithms (General Linear Model-GLM, General Additive Model-GAM, Random Forest-RF, Classification Tree Analysis-CTA, and MaxEnt) and the Ensemble model. Probability of presence ranges from 100 (red areas) to 0 (blue areas), while for the binary model it ranges from 1 (red areas) to 0 (white areas) (the threshold was defined optimizing TSS values from the Ensemble model, see main text for details).

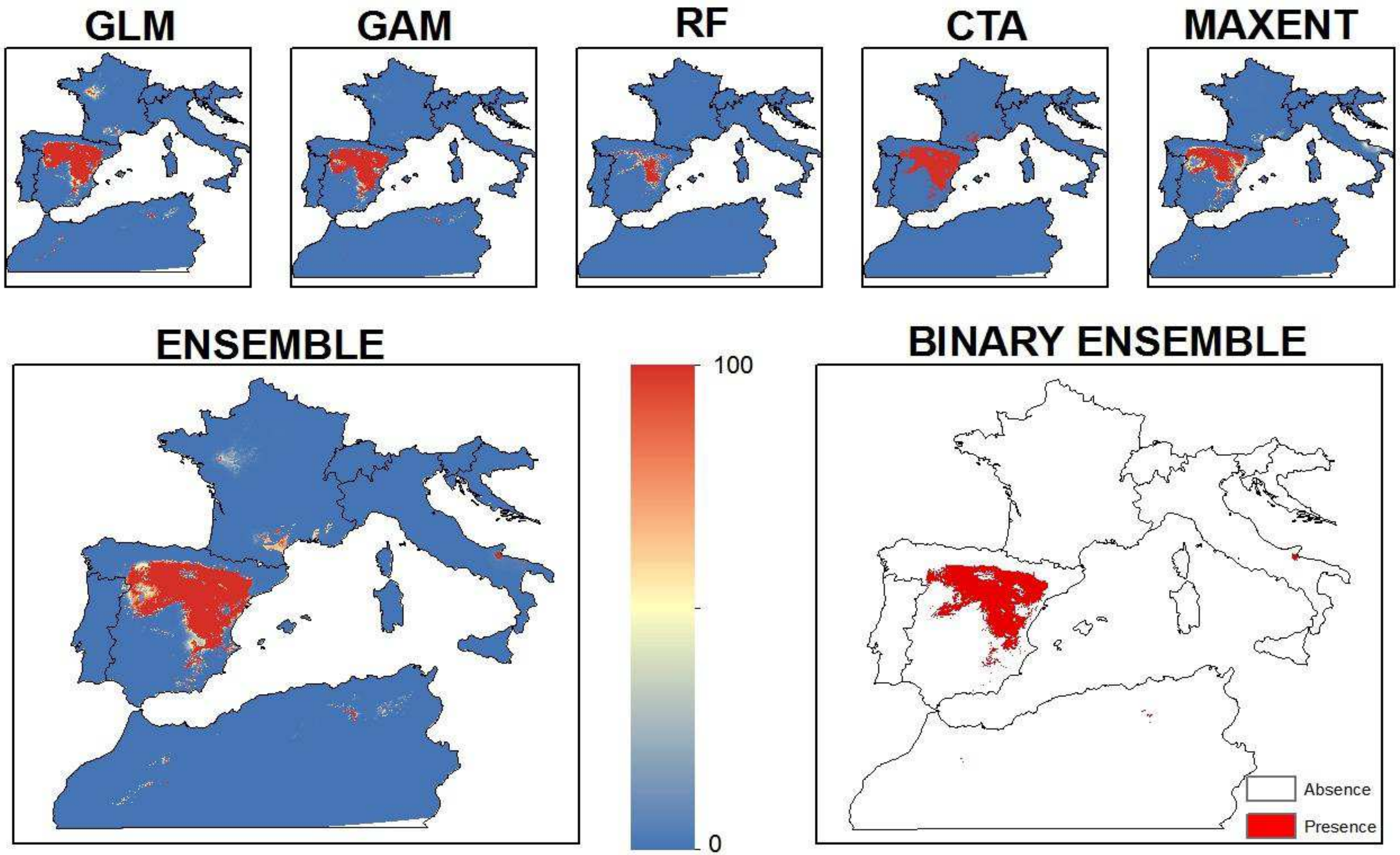


Figure S8: Geographic projection of the Eastern Spanish clade – G5 using five algorithms (General Linear Model-GLM, General Additive Model-GAM, Random Forest-RF, Classification Tree Analysis-CTA, and MaxEnt) and the Ensemble model. Probability of presence ranges from 100 (red areas) to 0 (blue areas), while for the binary model it ranges from 1 (red areas) to 0 (white areas) (the threshold was defined optimizing TSS values from the Ensemble model, see main text for details).

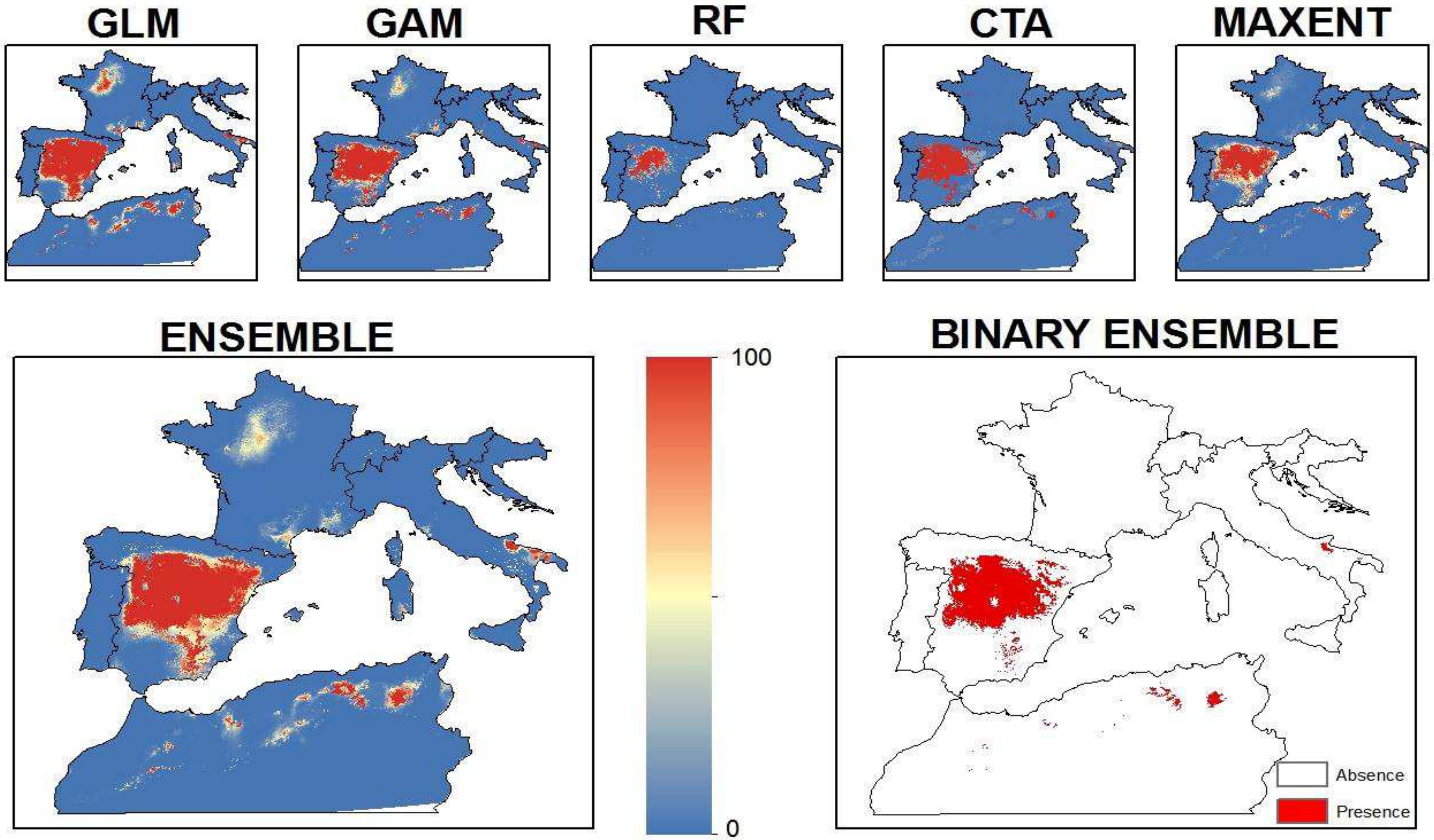


Figure S9: Geographic projection of the Central Spanish clade – G6 using five algorithms (General Linear Model-GLM, General Additive Model-GAM, Random Forest-RF, Classification Tree Analysis-CTA, and MaxEnt) and the Ensemble model. Probability of presence ranges from 100 (red areas) to 0 (blue areas), while for the binary model it ranges from 1 (red areas) to 0 (white areas) (the threshold was defined optimizing TSS values from the Ensemble model, see main text for details).

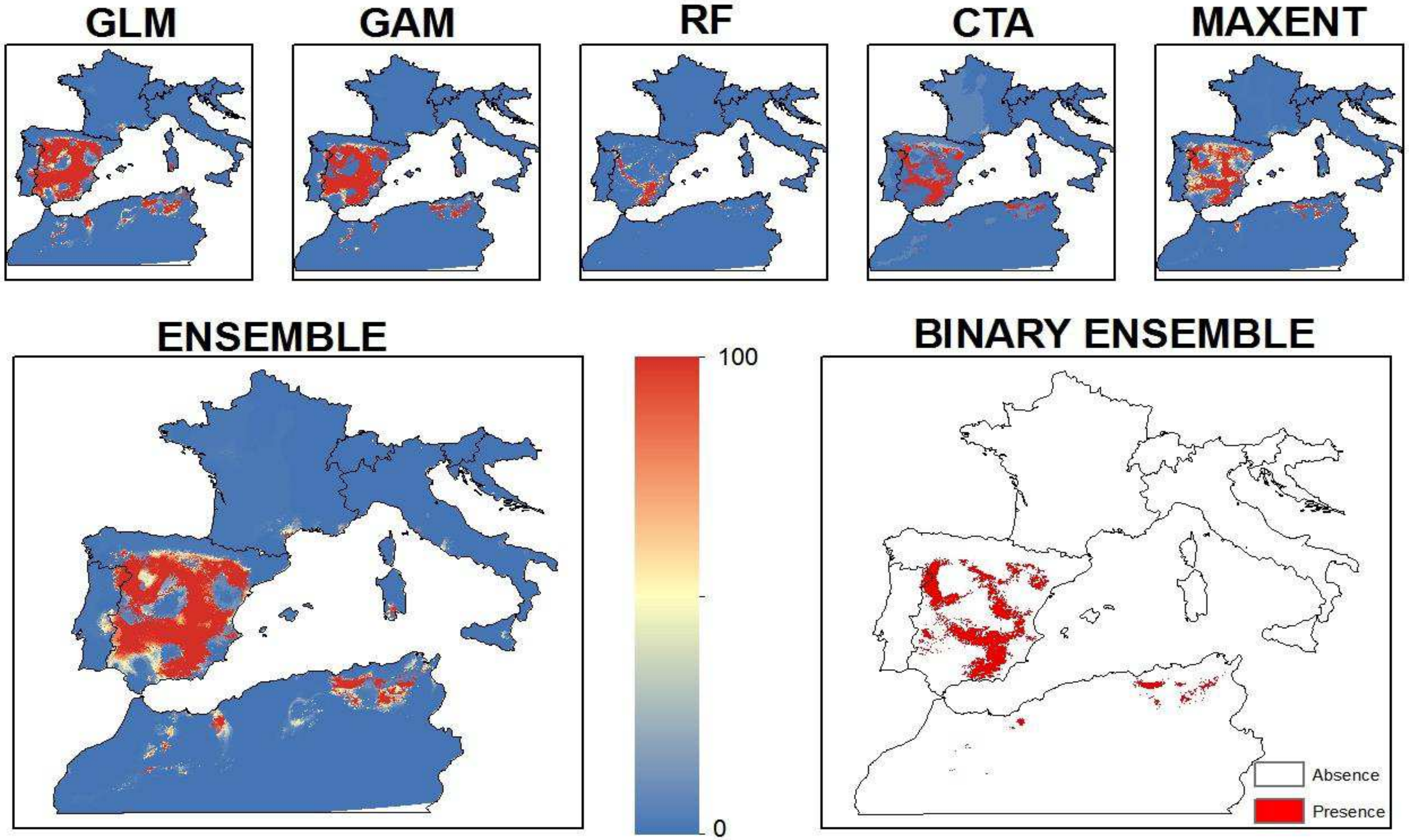


Figure S10: Geographic projection of the Southern Spanish clade – G7 using five algorithms (General Linear Model-GLM, General Additive Model-GAM, Random Forest-RF, Classification Tree Analysis-CTA, and MaxEnt) and the Ensemble model. Probability of presence ranges from 100 (red areas) to 0 (blue areas), while for the binary model it ranges from 1 (red areas) to 0 (white areas) (the threshold was defined optimizing TSS values from the Ensemble model, see main text for details).

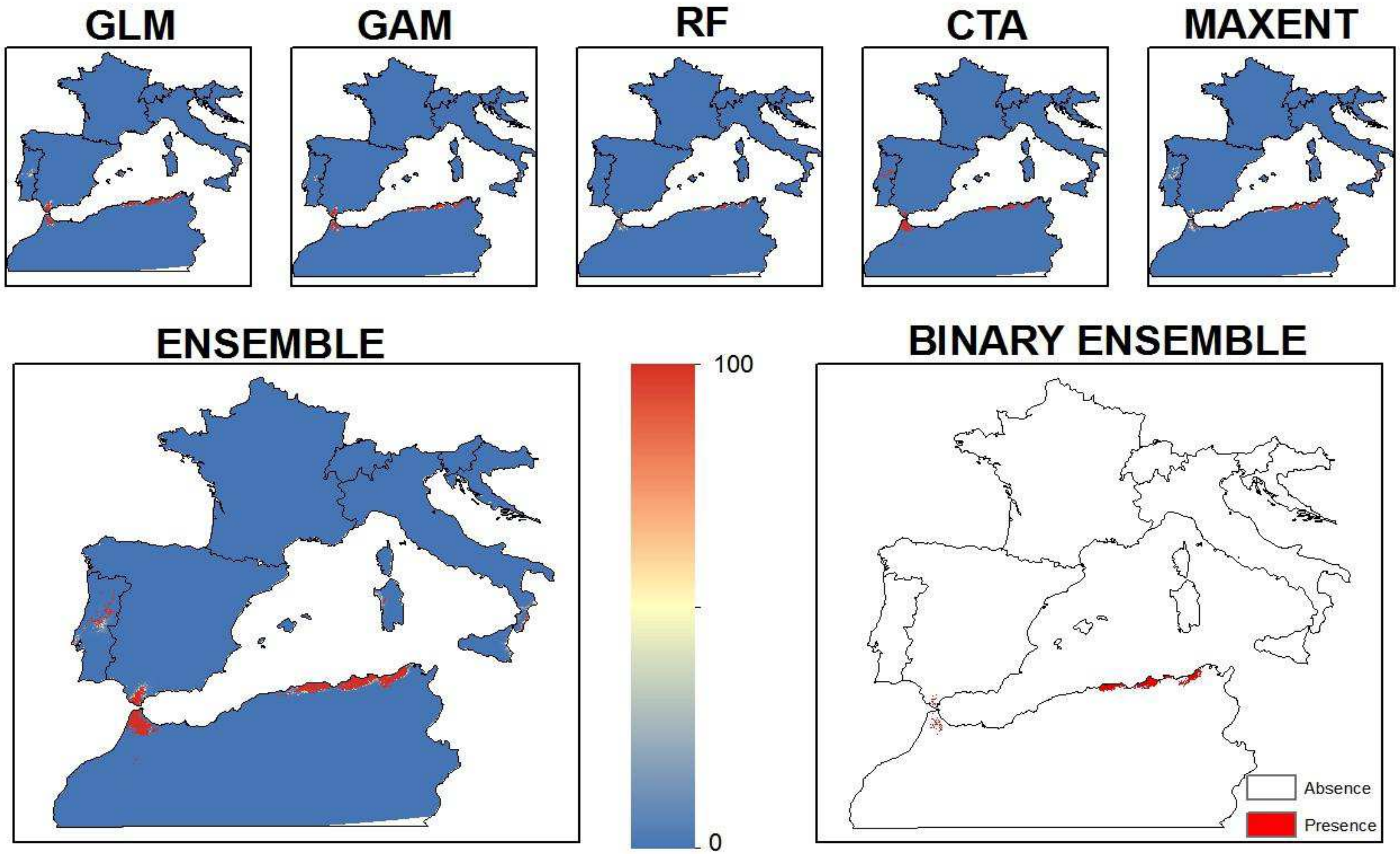


Figure S11: Geographic projection of the Tunisian clade – G8 using five algorithms (General Linear Model-GLM, General Additive Model-GAM, Random Forest-RF, Classification Tree Analysis-CTA, and MaxEnt) and the Ensemble model. Probability of presence ranges from 100 (red areas) to 0 (blue areas), while for the binary model it ranges from 1 (red areas) to 0 (white areas) (the threshold was defined optimizing TSS values from the Ensemble model, see main text for details).

Supporting information Chapter 2:

APPENDIX S1:

Details on *Pinus halepensis* Mill. genetic clades: Q membership values and description if the clades

APPENDIX S2:

Variable selection and evaluation scores obtained in species distribution models

APPENDIX S3:

A. Summary of conservation guidelines for *Pinus pinaster* Ait.
B. Summary of conservation guidelines for *Pinus halepensis* Mill.
C. Figures for *Pinus pinaster* Ait. clades (Fig. S1 – Fig. S8)
D. Figures for *Pinus halepensis* Mill. clades (Fig S9 – Fig S15)
E. Summary table of areas assigned to different conservation strategies (Table S6)

APPENDIX S4:

Dynamic Conservation Units: Classification within clades and Exposure assessment

APPENDIX S1

Details on *Pinus halepensis* Mill. genetic clades: Q membership values and description of the clades.

We detected seven different genetic clades by performing a Bayesian clustering analysis using STRUCTURE (Pritchard *et al.*, 2000) on a 294 Single Nucleotide Polymorphisms (SNPs) dataset (1325 individuals from 49 populations covering most of *Pinus halepensis* natural range); we estimated membership coefficients (Q*) for each individual within each clade (K); finally, we calculated definitive membership values (Q) for the 49 populations as a mean-over-population of individual Q* coefficients to each of the 7 gene pools defined when using Structure (see Table S1). Contrary to *Pinus pinaster* where the clades are spatially differentiated, we defined transition areas, occupied by more than one clade simultaneously, based on Q values as follows: (i) a population was classified as "pure" for one gene pool (K) if its Q values scored above 0.80 for that specific clade K; (ii) a population with Q values above 0.25 for one K was considered to have individuals belonging to that K (and thus that K was considered present in this population). Following this methodology, we obtained both pure populations and populations including up to two different genetic clades.

Below, we briefly describe the seven detected genetic clades namely Central and southern Spain clade (G1-hal), Balearic and southern France clade (G2-hal), Tunisian and northern Italian clade (G3-hal), Moroccan and southern Spain clade (G4-hal), Greek clade (G5-hal), Central and northern Spain clade (G6-hal) and Northern Spain and southern France clade (G7-hal). Some small areas of the distribution of the species (located in Algeria) could not be assigned to any clade due to insufficient sampling (see Fig. 2b in main text).

Brief description of the detected genetic clades

- ✓ Central and southern Spain clade (G1-hal): This clade coexists with G4-hal in southern Spain and with G6-hal in central Spain.

- ✓ Balearic and southern France clade (G2-hal): This clade is the only gene pool present in the Balearic Islands. Nevertheless, we can also find G2-hal in southern France where it coexists with G7-hal.
- ✓ Eastern Mediterranean clade (G3-hal): This clade is largely distributed in northern Africa (Tunisia) where it does not coexist with any other clade. We can also find it in Israel and in the Balkans, as well as in the Italian Peninsula where it shares territory with G5-hal.
- ✓ Moroccan and southern Spain clade (G4-hal): This clade represents the Moroccan distribution of *Pinus halepensis* and it also coexists with G1-hal in southern Spain.
- ✓ Greek clade (G5-hal): This clade is widely distributed in the eastern side of the Mediterranean Basin (Italy, Israel and the Balkans) along with G3-hal. Nevertheless, it is the only clade present in the Greek distribution of the species but for four marginal populations in the eastern-most distribution of the species.
- ✓ Central and northern Spain clade (G6-hal): This clade is distributed in central Spain (along with G1-hal) and in northern Spain along with G7-hal.
- ✓ Northern Spain and southern France clade (G7-hal): This clade co-occurs with G6-hal in the northern Spain and with G2-hal in southern France.

Table S1: Estimated membership coefficients (*Q*) for each population. Q values above 0.80 for one clade (K) correspond with pure populations for that specific K. Q coefficients above 0.25 for one K indicate presence of that K into the population. Colours correspond to overall genetic clusters defined for the study.

Population name	Long	Lat	K1	K2	K3	K4	K5	K6	K7
Amfilohia	21.28	38.88	0.00	0.01	0.01	0.01	0.97	0.00	0.00
North Eubea	23.18	38.58	0.00	0.00	0.01	0.00	0.97	0.00	0.00
Elea	21.53	37.77	0.00	0.00	0.00	0.00	0.98	0.00	0.00
Zaouia Ifrane	-5.14	33.57	0.00	0.00	0.01	0.97	0.00	0.00	0.00
Kassandra	23.88	40.09	0.01	0.01	0.34	0.02	0.62	0.01	0.01
Imperia	8.05	43.90	0.01	0.33	0.02	0.02	0.00	0.02	0.61
Shaharia	34.83	31.60	0.01	0.01	0.96	0.01	0.01	0.00	0.00
Litorale Tarantino	17.12	40.62	0.01	0.01	0.54	0.02	0.40	0.01	0.01
Thala	8.65	35.57	0.02	0.02	0.87	0.05	0.02	0.02	0.02
Aures Beni Melloul	6.83	35.17	0.02	0.01	0.83	0.07	0.01	0.03	0.04
Quercianella	10.34	43.49	0.02	0.03	0.58	0.01	0.30	0.02	0.04
Tabarka	9.08	36.51	0.02	0.01	0.89	0.05	0.00	0.01	0.01
Alcotx	4.17	39.97	0.02	0.70	0.01	0.01	0.00	0.02	0.24
Carlo Forte	8.18	39.08	0.02	0.75	0.16	0.01	0.00	0.04	0.01
Santanyi	3.05	39.28	0.03	0.83	0.03	0.02	0.00	0.05	0.04
Otricoli	12.38	42.24	0.04	0.06	0.61	0.01	0.24	0.03	0.01
Gargano Marzini	15.94	41.90	0.05	0.01	0.51	0.01	0.35	0.02	0.05
Garzano Monte Pucci	15.86	41.55	0.05	0.01	0.49	0.01	0.37	0.01	0.06
Nat	35.03	32.72	0.05	0.00	0.52	0.01	0.37	0.01	0.04
Palma de Mallorca	2.94	39.15	0.09	0.72	0.00	0.01	0.00	0.15	0.03
Alcudia	3.17	39.87	0.10	0.75	0.01	0.01	0.00	0.05	0.07
Zuera	-0.92	41.92	0.11	0.13	0.02	0.02	0.00	0.23	0.50
Atalix	4.05	39.92	0.11	0.71	0.01	0.02	0.00	0.06	0.09
Cabanellas	2.78	42.25	0.12	0.06	0.01	0.01	0.00	0.13	0.65
Tivissa	0.76	41.06	0.16	0.08	0.01	0.01	0.00	0.26	0.47
Serra d'Irta	0.32	40.35	0.18	0.05	0.02	0.03	0.00	0.30	0.42
Benicassim	0.03	40.08	0.24	0.02	0.01	0.00	0.00	0.33	0.39
Tuéjar	-1.16	39.82	0.28	0.03	0.01	0.01	0.00	0.62	0.05
Alhama de Murcia	-1.53	37.86	0.29	0.03	0.02	0.07	0.00	0.55	0.03
Benamaurel	-2.74	37.70	0.29	0.01	0.01	0.13	0.00	0.53	0.03
Frigiliana	-3.92	36.82	0.29	0.02	0.01	0.32	0.00	0.20	0.16
Carratraca	-4.83	36.84	0.29	0.01	0.00	0.39	0.00	0.12	0.19
Tibi	-0.65	38.52	0.30	0.04	0.02	0.06	0.00	0.55	0.03
Serra d'Irta	0.32	40.35	0.32	0.04	0.01	0.02	0.00	0.14	0.48
Bicorp	-0.86	39.10	0.33	0.07	0.02	0.03	0.00	0.52	0.03
Serra Calderona	-0.48	39.74	0.34	0.03	0.05	0.04	0.01	0.45	0.09
Alcantud	-2.31	40.56	0.35	0.04	0.01	0.01	0.00	0.52	0.06
Colmenar de Oreja	-3.33	40.09	0.35	0.03	0.01	0.01	0.00	0.58	0.02
Santiago de la Espada	-2.47	38.23	0.37	0.04	0.01	0.07	0.00	0.46	0.04
Serra Calderona	-0.48	39.74	0.38	0.02	0.01	0.02	0.00	0.49	0.08
Villajoyosa	-0.30	38.50	0.39	0.03	0.02	0.03	0.00	0.49	0.04
Monovar	-0.96	38.39	0.40	0.05	0.02	0.04	0.01	0.45	0.03
Montan	-0.59	40.05	0.41	0.08	0.01	0.02	0.00	0.39	0.09
Cabanes	0.04	40.10	0.42	0.03	0.01	0.00	0.00	0.13	0.41
Sinarcas	-1.20	39.80	0.43	0.05	0.01	0.01	0.00	0.46	0.04
Titaguas	-1.30	39.89	0.43	0.09	0.01	0.01	0.00	0.40	0.05
Cabanes	0.04	40.10	0.47	0.02	0.02	0.01	0.00	0.18	0.30
Alzira	-0.39	39.12	0.50	0.03	0.02	0.02	0.00	0.38	0.05
Eslida	-0.29	39.87	0.50	0.05	0.01	0.01	0.00	0.32	0.11

References

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

APPENDIX S2

Variable selection and evaluation scores obtained in the Species Distribution Models

Variable selection

We used the 19 bioclimatic variables available in WORLDCLIM (Hijmans *et al.*, 2005) representative of the period between the years 1950-2000 for the analysis.

In order to avoid multicollinearity effects, we retained variables with Pearson correlations lower than 0.75 as the use of simple methods based on rules of thumb have proved to be as effective as more complicated ones (Dormann *et al.,* 2013). This threshold value is a bit less restrictive than the common value of 0.70 but more flexible thresholds have also been used in literature (see Elith *et al.,* 2006 for an example). Among highly correlated variables we kept the one with highest explained deviance scores (D^2) when individually fitted in a Generalized Linear Model (GLM; McCullagh & Nelder, 1989). The similar patterns detected in D^2 values in *Pinus pinaster* and *Pinus halepensis* scores enabled the selection of the same set of variables for both species, which also eased the subsequent interpretation of the results. We discarded BIO15 (due to its very low D^2 scores: 0.05 and 0.11 for *P. pinaster* and *P. halepensis* respectively) and BIO8 (as the steep gradient shown by this variable may lead to artefacts in the SDM output maps, because very often two adjacent cells are characterized by extremely different values within the study area for no obvious reason). Finally we performed a Variance Inflation Factor (VIF) test and removed BIO1 to ensure that all VIF values were below a threshold value of 10.

The final set of relevant weakly correlated variables was BIO4 (Temperature Seasonality), BIO11 (Mean Temperature of Coldest Quarter), BIO12 (Annual Precipitation) and BIO18 (Precipitation of Warmest Quarter).

Future climate projections

We used all Global Climate Models (GCMs) that were simultaneously available for the scenarios of representative concentration pathways (RCP) 2.6, 4.5 and 8.5. This included 14 GCMs: BCC-CSM1-1 (BC), CCSM4 (CC), CNRM-CM5 (CN), GFDL-CM3 (GF), GISS-E2-R (GS), HadGEM2-AO (HD), HadGEM2-ES (HE), IPSL-CM5A-LR (IP), MIROC-ESM-CHEM (MI), MIROC-ESM (MR), MIROC5 (MC), MPI-ESM-LR (MP), MRI-CGCM3 (MG), NorESM1-M (NO).

Evaluation scores obtained in the Species Distribution Models

Table S2: Assessment of *Pinus pinaster* Ait. model performance by means of True Skill Statistic (TSS) and Area Under the Curve (AUC). Values are displayed separately according to genetic clade level (G1-pin to G8-pin) and also to the five different algorithms employed: General Linear Model (GLM), General Additive Model (GAM), Random Forest (RF), Classification Tree Analysis (CTA) and MaxEnt.

Model	Algorithm	AUC	TSS
G1-pin	GLM	0.996	0.964
	GAM	0.997	0.981
	RF	0.999	0.990
	CTA	0.994	0.976
	MaxEnt	0.997	0.977
G2-pin	GLM	0.982	0.871
	GAM	0.985	0.948
	RF	1.000	0.983
	CTA	0.985	0.967
	MaxEnt	0.988	0.914
G3-pin	GLM	0.999	0.996
	GAM	1.000	0.997
	RF	1.000	0.997
	CTA	0.997	0.994
	MaxEnt	0.999	0.990
G4-pin	GLM	0.804	0.710
	GAM	0.986	0.943
	RF	1.000	0.960
	CTA	0.993	0.972
	MaxEnt	0.992	0.931
G5-pin	GLM	0.996	0.974
	GAM	0.998	0.989
	RF	1.000	0.975
	CTA	0.993	0.977
	MaxEnt	0.998	0.972
G6-pin	GLM	0.986	0.933
	GAM	0.993	0.970
	RF	1.000	0.982
	CTA	0.991	0.977
	MaxEnt	0.993	0.945
G7-pin	GLM	0.973	0.898
	GAM	0.989	0.951
	RF	0.999	0.943
	CTA	0.982	0.963
	MaxEnt	0.989	0.917
G8-pin	GLM	1.000	0.997
	GAM	1.000	0.998
	RF	1.000	0.986
	CTA	0.994	0.988

	MaxEnt	0.999	0.989

Table S3: Assessment of *Pinus halepensis* Mill. model performance by means of True Skill Statistic (TSS) and Area Under the Curve (AUC). Values are displayed separately according to genetic clade level (G1-hal to G7-hal) and also to the five different algorithms employed: General Linear Model (GLM), General Additive Model (GAM), Random Forest (RF), Classification Tree Analysis (CTA) and MaxEnt.

Model	**Algorithm**	**AUC**	**TSS**
G1-hal	GLM	0.970	0.870
	GAM	0.889	0.889
	RF	0.999	0.971
	CTA	0.969	0.917
	MaxEnt	0.977	0.869
G2-hal	GLM	0.978	0.915
	GAM	0.990	0.962
	RF	0.999	0.973
	CTA	0.988	0.967
	MaxEnt	0.990	0.895
G3-hal	GLM	0.989	0.916
	GAM	0.991	0.942
	RF	0.999	0.987
	CTA	0.985	0.951
	MaxEnt	0.991	0.923
G4-hal	GLM	0.987	0.935
	GAM	0.994	0.936
	RF	0.999	0.969
	CTA	0.987	0.953
	MaxEnt	0.992	0.919
G5-hal	GLM	0.983	0.959
	GAM	0.988	0.959
	RF	0.998	0.928
	CTA	0.981	0.955
	MaxEnt	0.990	0.939
G6-hal	GLM	0.973	0.835
	GAM	0.979	0.883
	RF	0.998	0.977
	CTA	0.979	0.926
	MaxEnt	0.982	0.872
G7-hal	GLM	0.980	0.914
	GAM	0.984	0.929
	RF	0.999	0.971
	CTA	0.976	0.940
	MaxEnt	0.987	0.907

Table S4: True Skill Statistic (TSS), Cutoff, Sensitivity and Specificity scores of the Ensemble model (average of the five algorithms), for the models of the clades of *Pinus pinaster* Ait. (G1-pin to G8-pin).

Model	TSS	Cutoff	Sensitivity	Specificity
G1-pin	0.986	552	99.503	99.096
G2-pin	0.967	479	98.855	97.979
G3-pin	0.997	415	99.955	99.779
G4-pin	0.968	584	97.624	99.187
G5-pin	0.987	466	99.591	99.113
G6-pin	0.981	536	99.154	98.989
G7-pin	0.946	589	96.201	98.402
G8-pin	0.996	441	99.704	99.904

Table S5: True Skill Statistic (TSS), Cutoff, Sensitivity and Specificity scores of the Ensemble model (average of the five algorithms), for the models of the clades of *Pinus halepensis* Mill. (G1-hal to G7-hal).

Model	TSS	Cutoff	Sensitivity	Specificity
G1-hal	0.927	481	97.448	94.350
G2-hal	0.968	496	98.542	98.412
G3-hal	0.959	595	97.453	98.429
G4-hal	0.969	508	98.626	98.265
G5-hal	0.952	632	96.492	98.719
G6-hal	0.927	470	97.172	95.576
G7-hal	0.945	601	97.069	97.461

References

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., Mcclean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D. & Lautenbach, S. (2013) Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 027–046.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-pereira, R., Schapire, R.E., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species ' distributions from occurrence data. *Ecography*, **29**, 129–151.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

McCullagh, P. & Nelder, J.A. (1989) *Generalized linear models*, (ed. by C. and Hall).

APPENDIX S3

A: Summary of conservation guidelines for *Pinus pinaster* Ait.

G1-pin: In general, the conservation of G1-pin can be easily guaranteed as areas for *in situ* conservation are largely available. There are also areas for translocation in northern Spain for current actions and northern France for midterm performance although they should not be necessary.

G2-pin: G2-pin is very likely to withstand climate change as only ca. 15 % of its distribution is largely exposed. There are *in situ* conservation areas available where in principle no monitoring is needed. There are also large areas proposed for translocation in case it would be necessary both currently and in the midterm (in Italy, southern France and Turkey for the former and northern France for the later).

G3-pin: From the currently occupied territory of G3-pin, almost 50% is highly exposed and thus designated for *ex situ* conservation. There are not areas proposed for *in situ* conservation without monitoring. In parallel, there are new areas very likely to become suitable for the clade in the future in United Kingdom (classified as likely suitable in the future) which provides the possibility of translocation in the midterm.

G4-pin: Approximately 70 % of the current distribution of this gene pool is classified as largely exposed. *In situ* conservation areas are available but monitoring is generally needed. In case translocation wants to be tested, there are a set of locations provided (Morocco for current actions and southern Spain for midterm performance).

G5-pin: This clade is largely exposed (over 95 % of its current distribution is designated for *ex situ* conservation). The scarce locations proposed for *in situ* conservation require monitoring. *Ex situ* conservation is thus the recommended option for this clade. Despite that the proposed areas for translocation are majorly secondary options, there are also priority territories available for both current and midterm performance.

G6-pin: This clade is highly exposed: its current distribution range is fully designated for *ex situ* conservation (100 %). In addition there are no areas suitable for the clade outside its distribution and new areas will not become suitable in the future. Thus, translocation is not very likely to be successful and the conservation of this clade requires new alternatives such as conservation in germplasm banks.

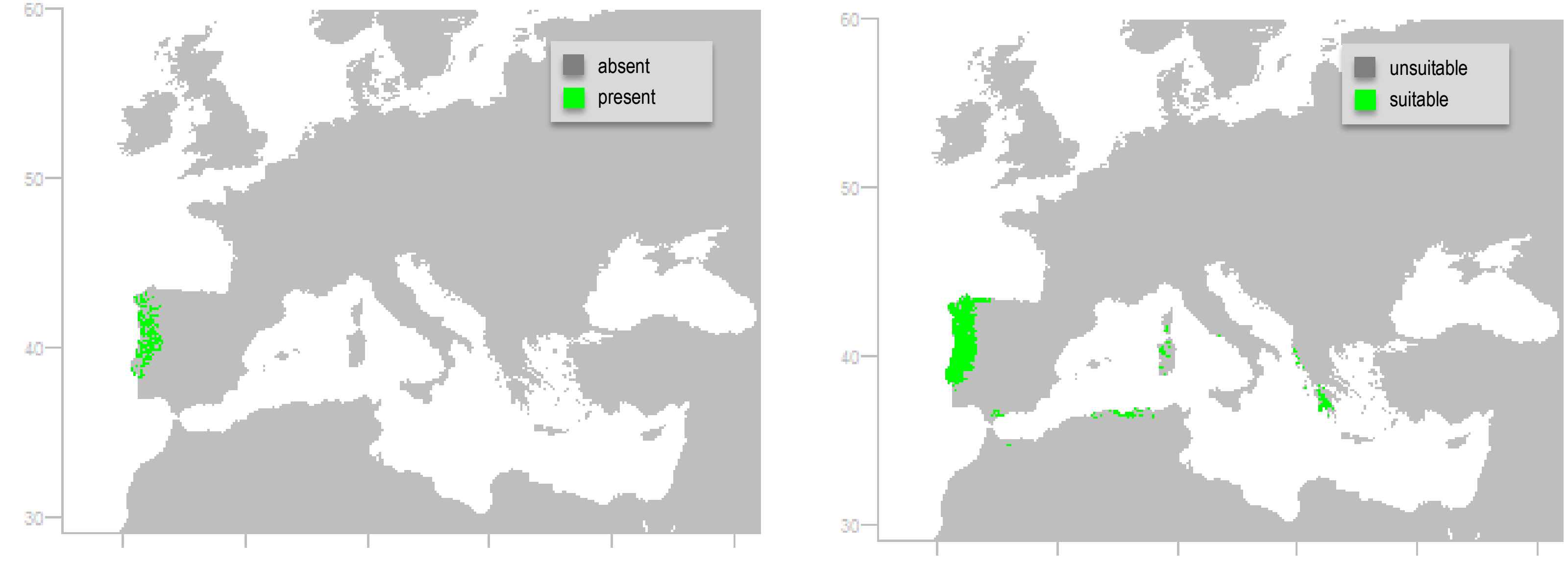
G7-pin: G7-pin is largely exposed as over 95 % of its territory is proposed for *ex situ* conservation. Thus, areas for *in situ* conservation are very scarce and require monitoring. As for *ex situ* conservation, assisted migration is possible as there are some priority assisted migration locations. Nevertheless, they are scarce and conservation in germplasm banks should not be dismissed.

G8-pin: Very low percent of the G8-pin clade is recommended for *ex situ* conservation and thus this is not a largely exposed clade. On the contrary, *in situ* conservation should be enough to guarantee the maintenance of this gene pool. Translocation is also possible (especially in the midterm) to some areas of Portugal, the Balkans and the Italian Peninsula.

B: Summary of conservation guidelines for *Pinus halepensis* Mill.

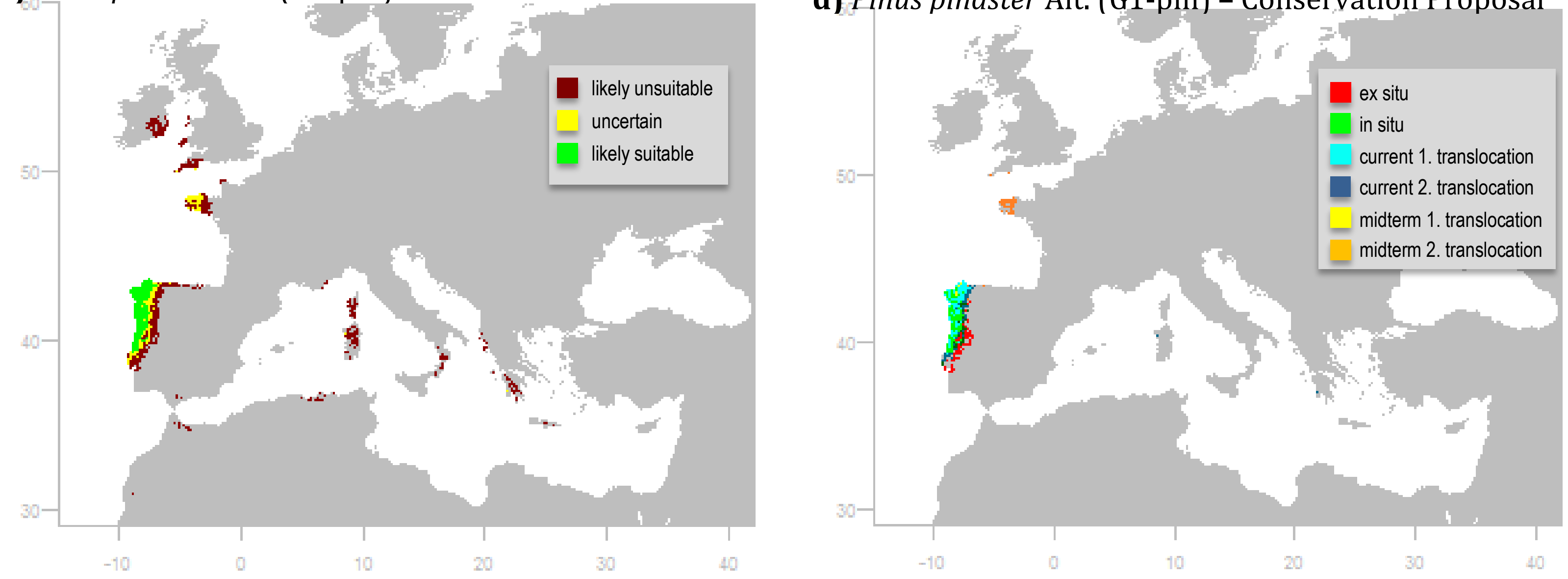
G1-hal:	In this clade, there are large areas proposed for *in situ* conservation and in most of them monitoring is not strictly necessary. Nevertheless, if *ex situ* conservation alternatives want to be explored there are translocation areas proposed (*e.g.* northern Spain for current actions).
G2-hal:	G2-hal has large areas proposed for *in situ* conservation and in most of them monitoring is not strictly necessary. There are also translocation areas available, although majorly to be applied in the midterm and with an intermediate probability of success (midterm secondary option translocation areas).
G3-hal:	G3-hal is likely to withstand climate change as only *ca.* 11 % of its currently occupied territory is under high risk of habitat loss, and almost all populations are capable of *in situ* conservation. In addition, the spatial extent of suitable territory is very likely to widen in the future enabling translocation with high probability of success.
G4-hal:	This gene pool is largely exposed (over 90 % of its currently occupied territory is recommended for *ex situ* conservation). Areas proposed for *in situ* conservation are scarce and require monitoring. On the contrary, translocation seems to be a good option as there are relatively large areas suited for it in southern Morocco and southern Spain.
G5-hal:	This clade is highly exposed as a large proportion of its distribution is designed for *ex situ* conservation (ca.80 %). Nevertheless, there are territories proposed for *in situ* conservation although mainly these need monitoring. There are also some current first option translocation areas very near the current distribution of the species (southern Greece).
G6-hal:	G6-hal is weakly exposed to climate change (only *ca.* 30 % of its currently occupied territory is under high risk of habitat loss). There are large *in situ* conservation areas and, if desired, translocation could also be performed (overall in Spain and Italy for current actions and northern France, Belgium and Netherlands for the midterm).
G7-hal:	Over 55 % of the distribution of G7-hal is recommended for *ex situ* conservation (thus considered highly exposed). Nevertheless, there are populations recommended for *in situ* conservation some of them with no need of monitoring. There are also first option translocation areas proposed (both for current and midterm performance) located in northern France in case *ex situ* conservation efforts want to be tested.

C.-Figures for *Pinus pinaster* Ait. clades



a) *Pinus pinaster* Ait. (G1-pin) - Current Distribution

b) *Pinus pinaster* Ait. (G1-pin) – Suitable Habitat Current Climate



c) *Pinus pinaster* Ait. (G1-pin) – Suitable Habitat 2050

d) *Pinus pinaster* Ait. (G1-pin) – Conservation Proposal

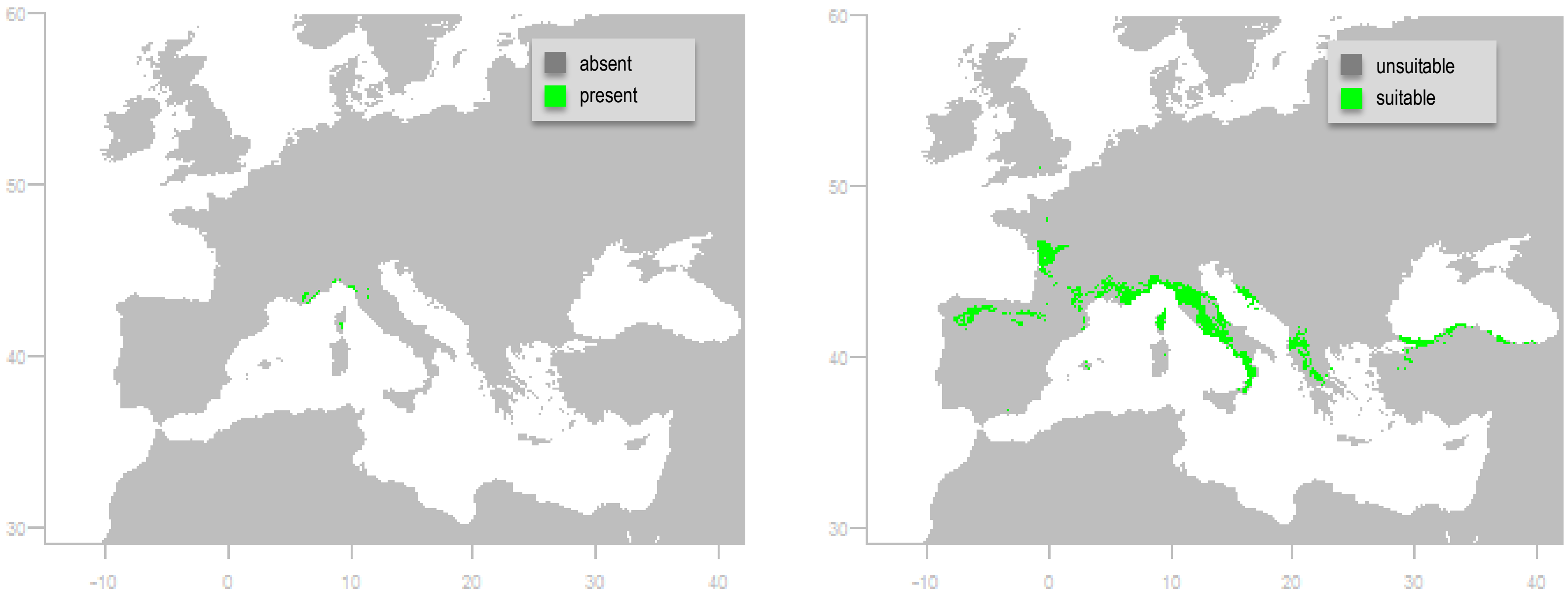
Figure S1:

a) Current distribution range of Atlantic Iberian Peninsula (G1-pin) of *Pinus pinaster* Ait.

b) Simulated habitat suitability under current climate for clade G1-pin.

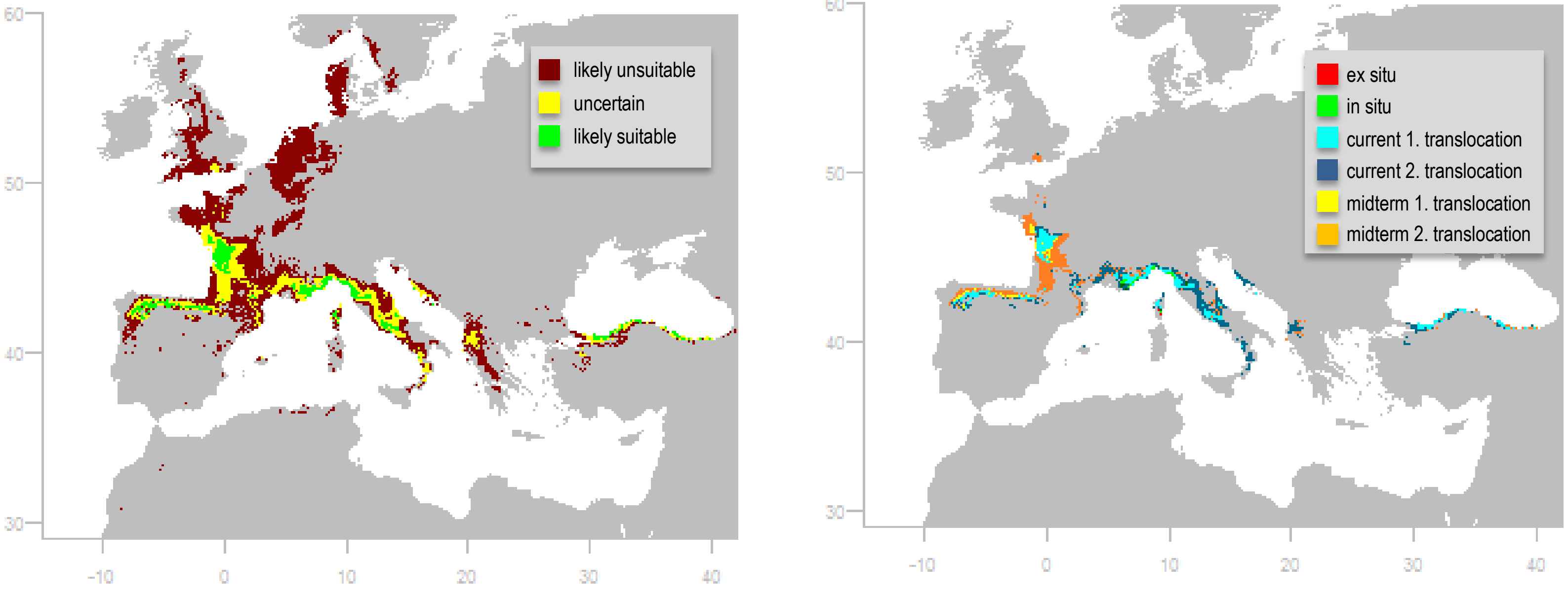
c) Classified future (2050) habitat suitability map of G1-pin.

d) Conservation strategies proposed



a) *Pinus pinaster* Ait. (G2-pin) - Current Distribution

b) *Pinus pinaster* Ait. (G2-pin) – Suitable Habitat Current Climate



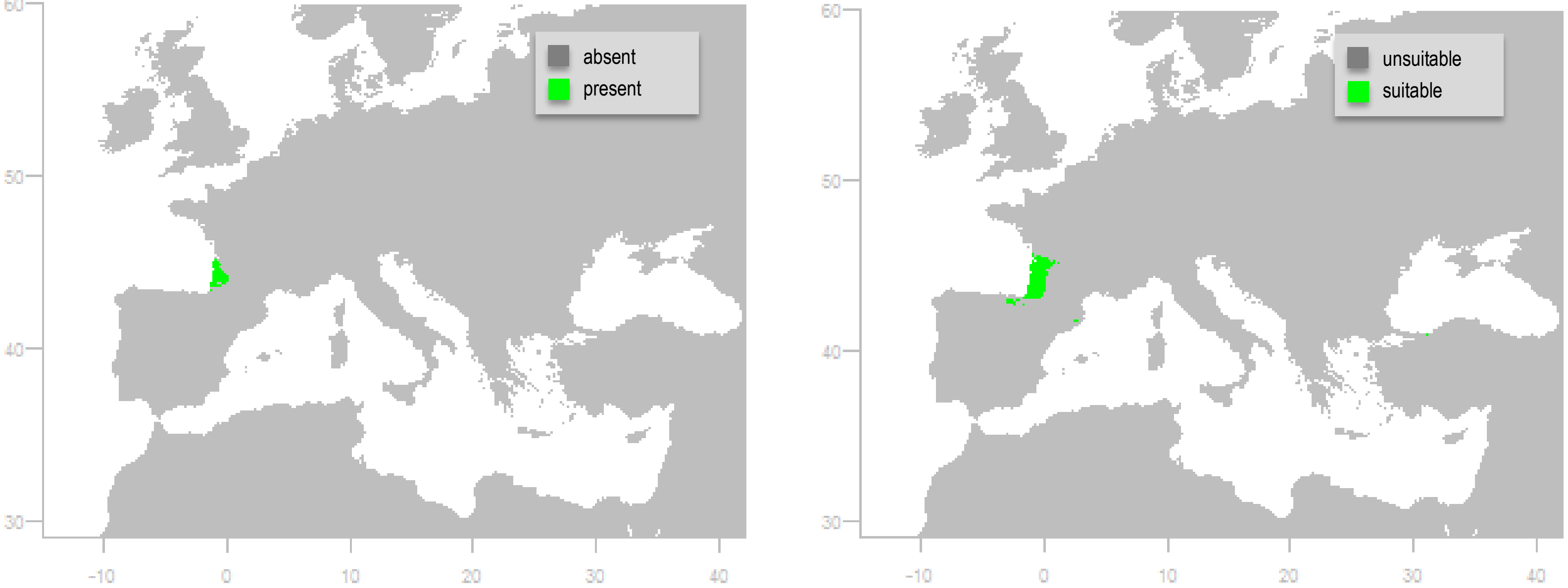
c) *Pinus pinaster* Ait. (G2-pin) – Suitable Habitat 2050

d) *Pinus pinaster* Ait. (G2-pin) – Conservation Proposal

Figure S2:
a) Current distribution range of the Eastern clade (G2-pin) of *Pinus pinaster* Ait.
b) Simulated habitat suitability under current climate for clade G2-pin.
c): Classified future (2050) habitat suitability map of G2-pin.
d): Conservation strategies proposed

a) *Pinus pinaster* Ait. (G3-pin) - Current Distribution

b) *Pinus pinaster* Ait. (G3-pin) – Suitable Habitat Current



c) *Pinus pinaster* Ait. (G3-pin) – Suitable Habitat 2050

d) *Pinus pinaster* Ait. (G3-pin) – Conservation Proposal

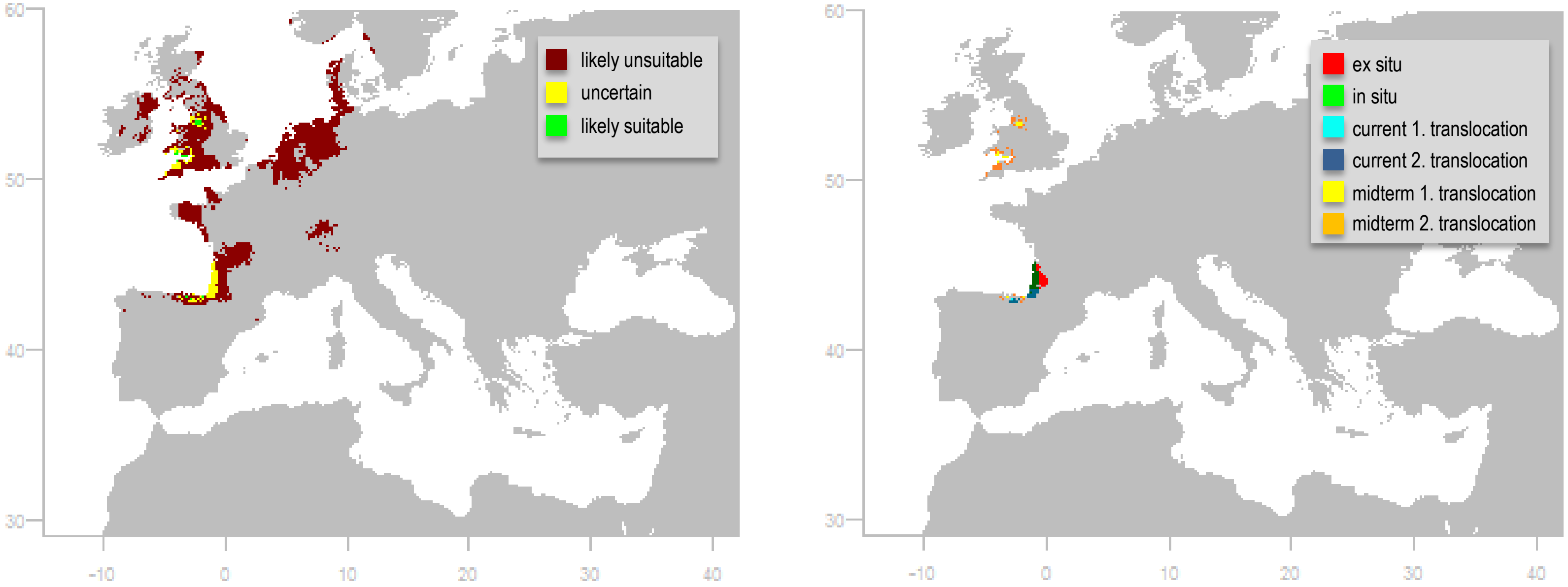
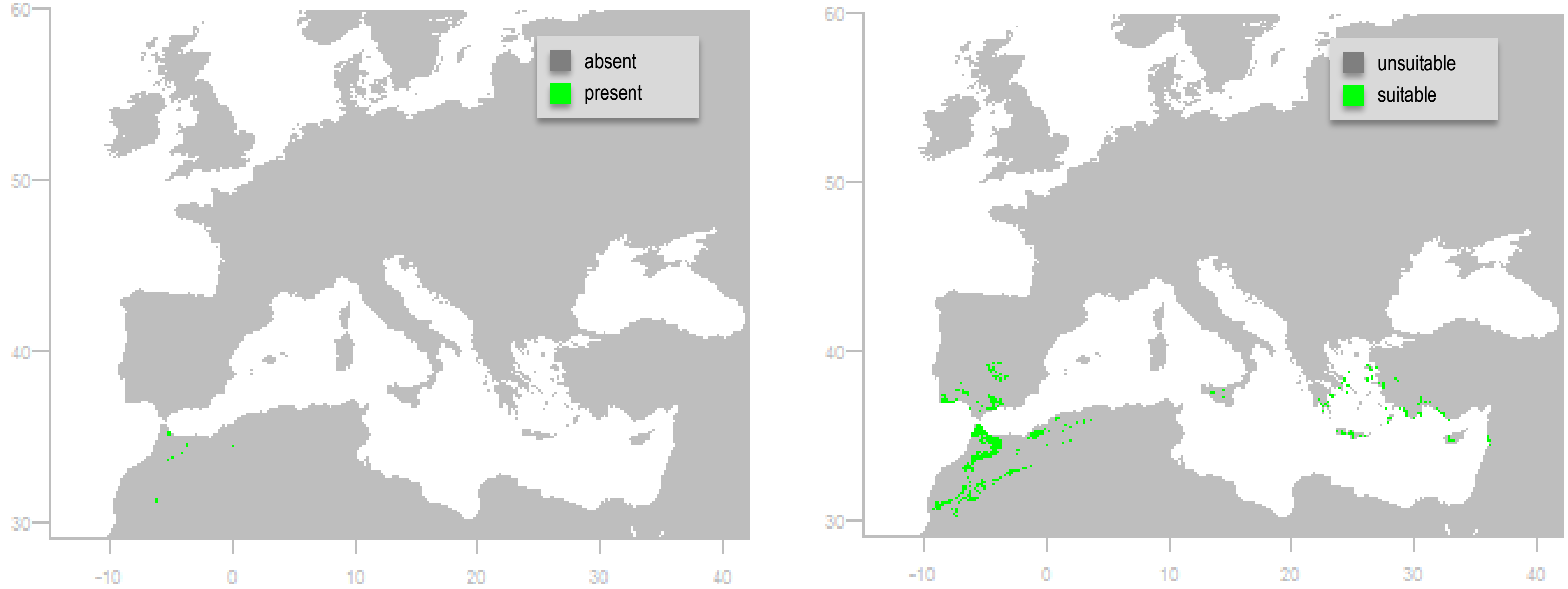


Figure S3:

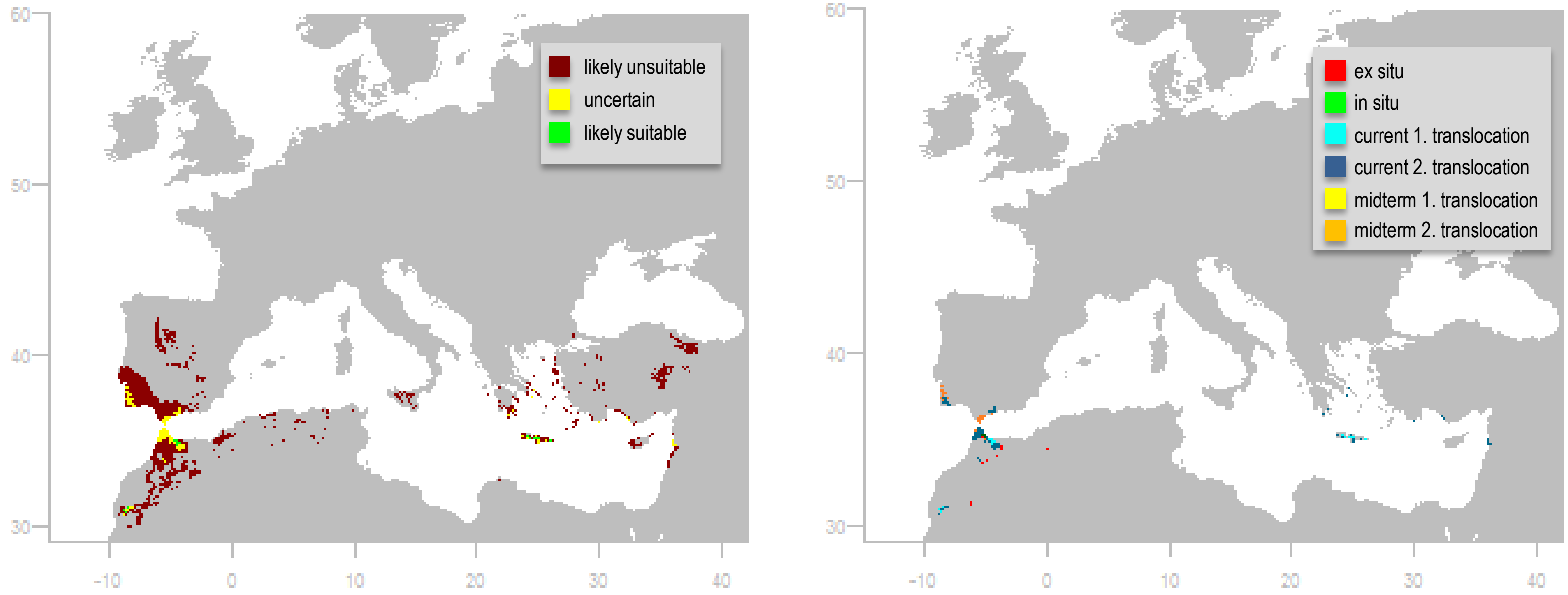
a) Current distribution range of the Atlantic France (G3-pin) of *Pinus pinaster* Ait.

b) Simulated habitat suitability under current climate for clade G3-pin. **c):** Classified future (2050) habitat suitability map of G3-pin.

d): Conservation strategies proposed for clade G3-pin.

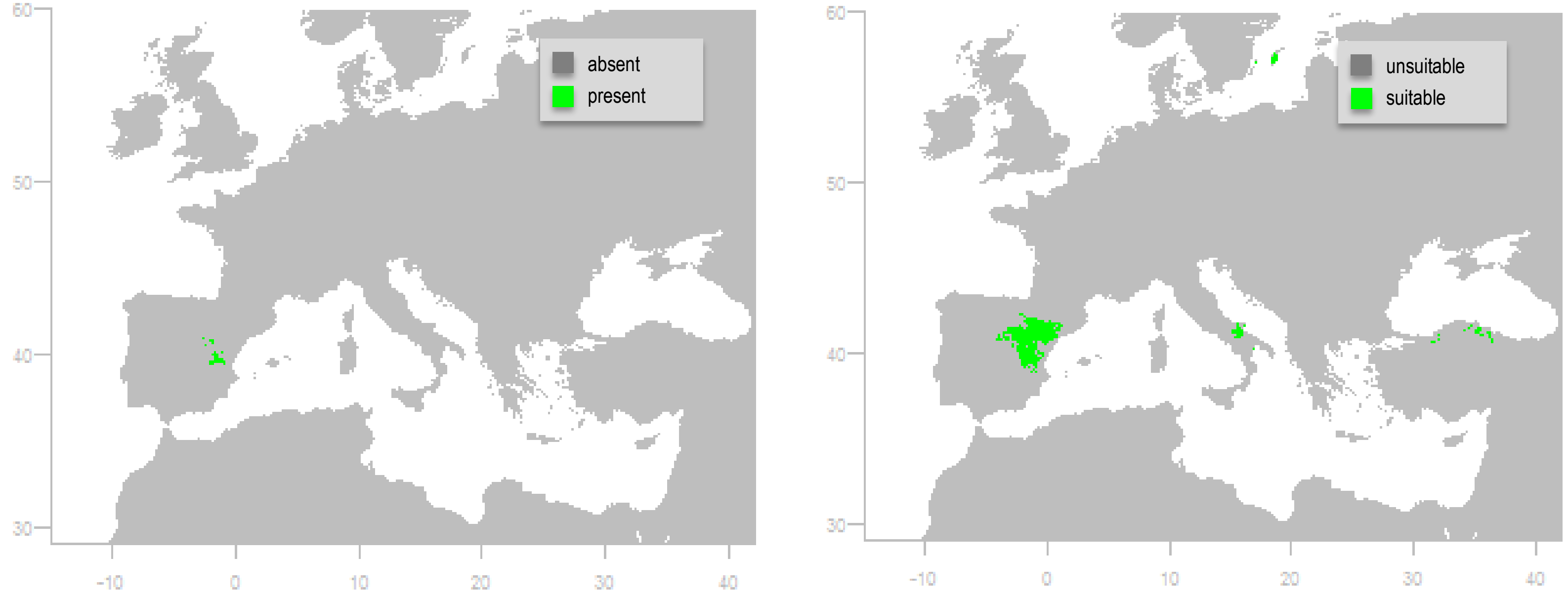

a) *Pinus pinaster* Ait. (G4-pin) - Current Distribution

b) *Pinus pinaster* Ait. (G4-pin) – Suitable Habitat Current Climate


c) *Pinus pinaster* Ait. (G4-pin) – Suitable Habitat 2050

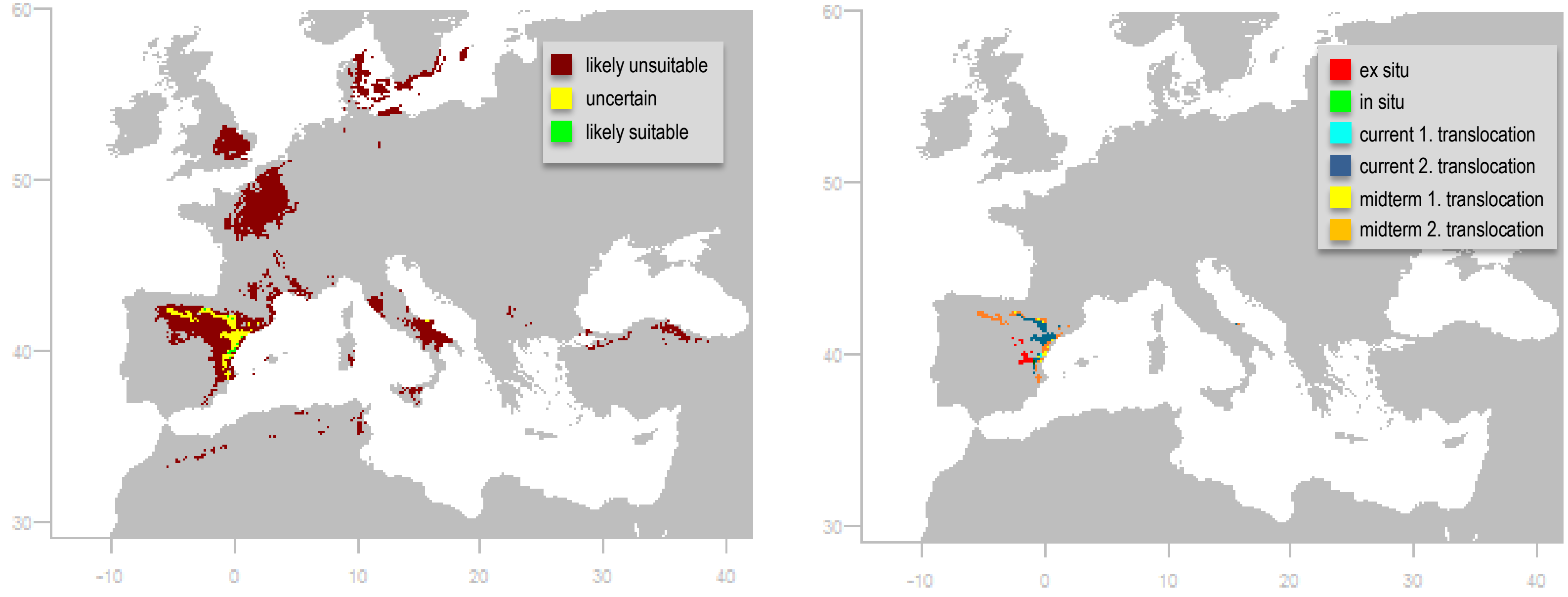
d) *Pinus pinaster* Ait. (G4-pin) – Conservation Proposal

Figure S4 :
a) Current distribution range of the Morocco clade (G4-pin) of *Pinus pinaster* Ait.
b) Simulated habitat suitability under current climate for clade G4-pin.
c): Classified future (2050) habitat suitability map of G4-pin.
d): Conservation strategies proposed for clade G4-pin.



a) *Pinus pinaster* Ait. (G5-pin) - Current Distribution

b) *Pinus pinaster* Ait. (G5-pin) – Suitable Habitat Current Climate

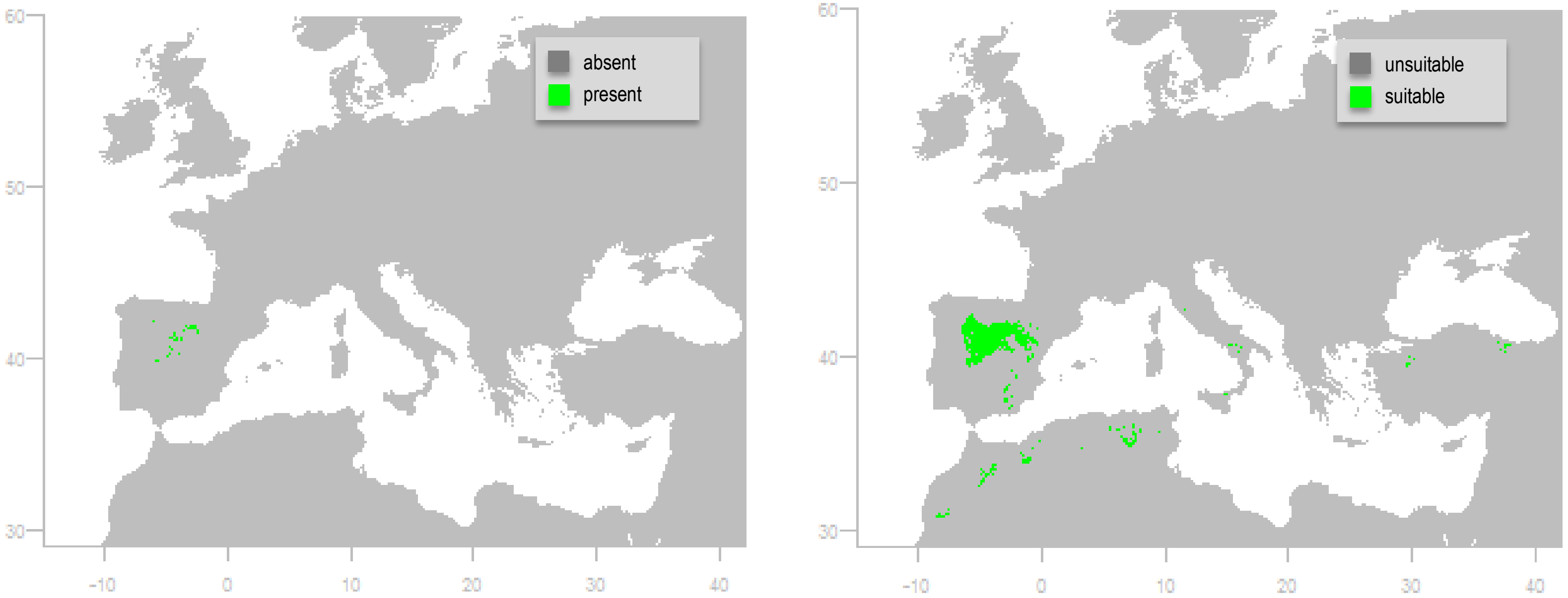


c) *Pinus pinaster* Ait. (G5-pin) – Suitable Habitat 2050

d) *Pinus pinaster* Ait. (G5-pin) – Conservation Proposal

Figure S5 :
a) Current distribution range of the Eastern Spain clade (G5-pin) of *Pinus pinaster* Ait.
b) Simulated habitat suitability under current climate for clade G5-pin. **c)** Classified future (2050) habitat suitability map of G5-pin.
d): Conservation strategies proposed for clade G5-pin.

a) *Pinus pinaster* Ait. (G6-pin) - Current Distribution **b)** *Pinus pinaster* Ait. (G6-pin) – Suitable Habitat Current Climate



c) *Pinus pinaster* Ait. (G6-pin) – Suitable Habitat 2050 **d)** *Pinus pinaster* Ait. (G6-pin) – Conservation Proposal

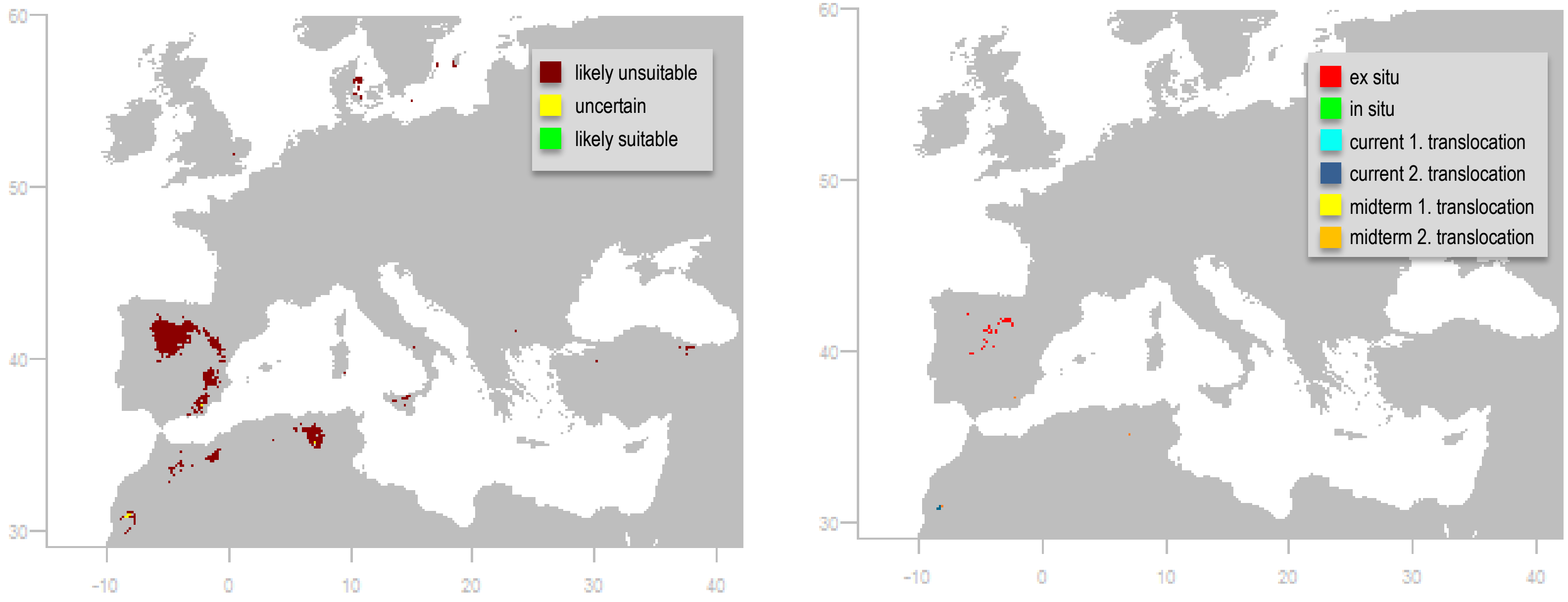


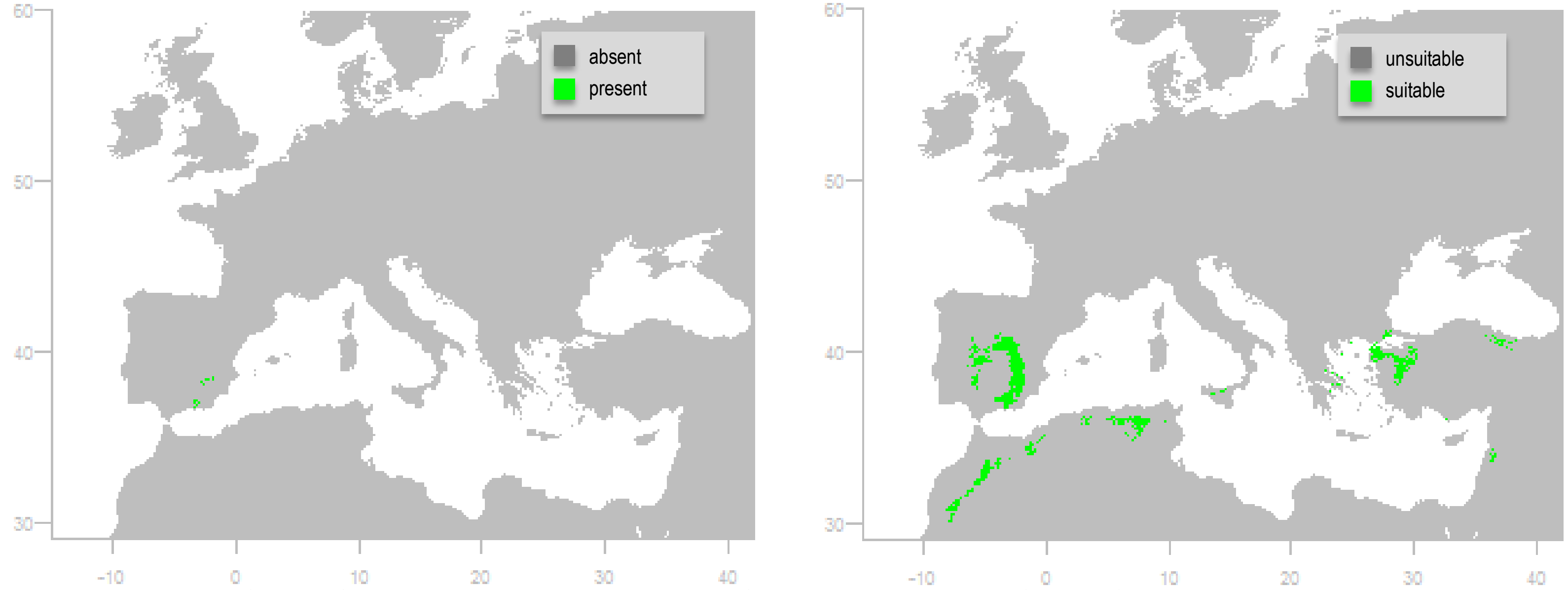
Figure S6 :

a) Current distribution range of the Central Spain clade (G6-pin) of *Pinus pinaster* Ait.

b) Simulated habitat suitability under current climate for clade G6-pin.

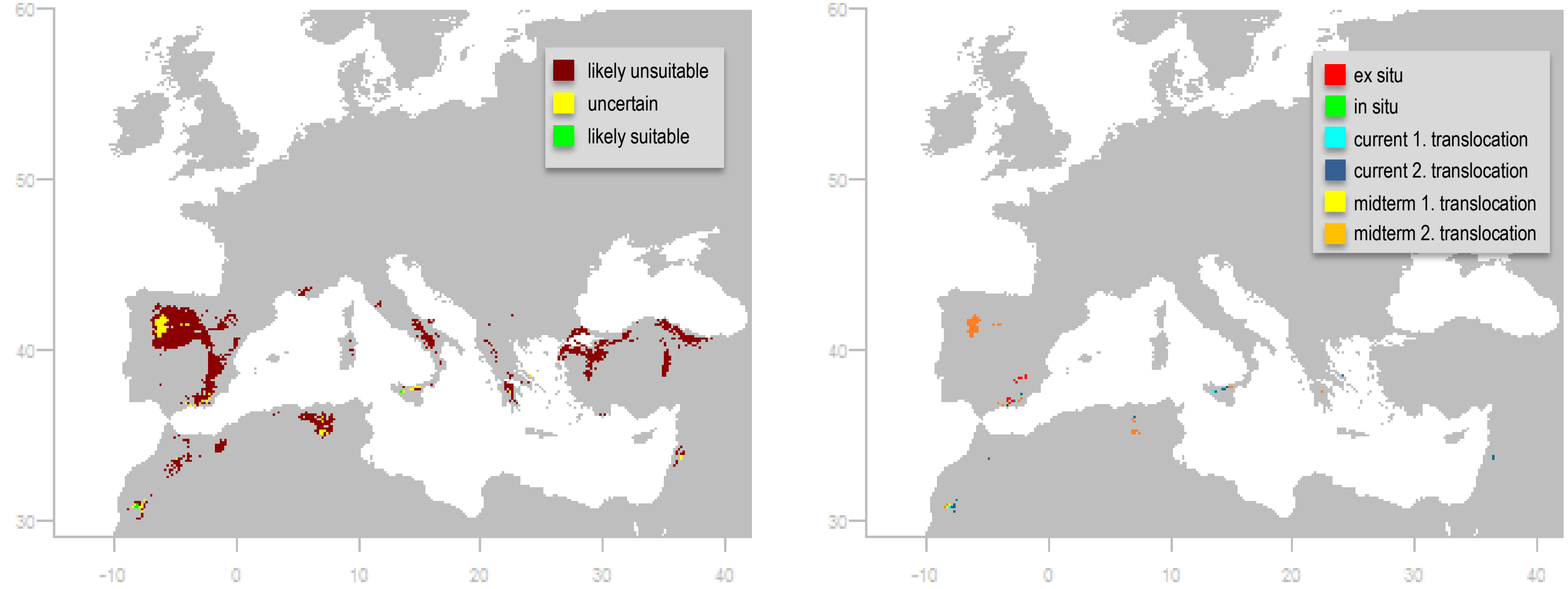
c) Classified future (2050) habitat suitability map of G6-pin.

d) Conservation strategies proposed for clade G6-pin.



a) *Pinus pinaster* Ait. (G7-pin) - Current Distribution

b) *Pinus pinaster* Ait. (G7-pin) – Suitable Habitat Current Climate

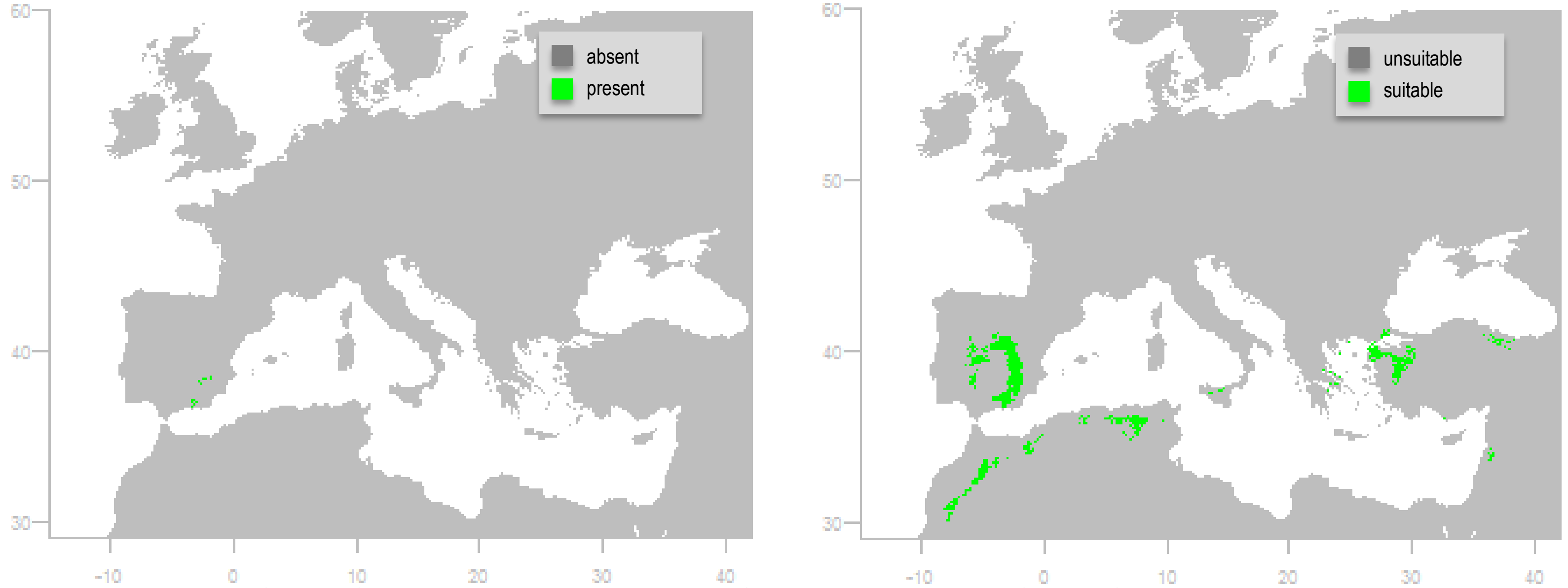


c) *Pinus pinaster* Ait. (G7-pin) – Suitable Habitat 2050

d) *Pinus pinaster* Ait. (G7-pin) – Conservation Proposal

Figure S7 :
a) Current distribution range of the Southern Spain clade (G7-pin) of *Pinus pinaster* Ait.
b) Simulated habitat suitability under current climate for clade G7-pin.
c) Classified future (2050) habitat suitability map of G7-pin.
d) Conservation strategies proposed for clade G7-pin.

a) *Pinus pinaster* Ait. (G8-pin) - Current Distribution **b)** *Pinus pinaster* Ait. (G8-pin) – Suitable Habitat Current Climate



c) *Pinus pinaster* Ait. (G8-pin) – Suitable Habitat 2050 **d)** *Pinus pinaster* Ait. (G8-pin) – Conservation Proposal

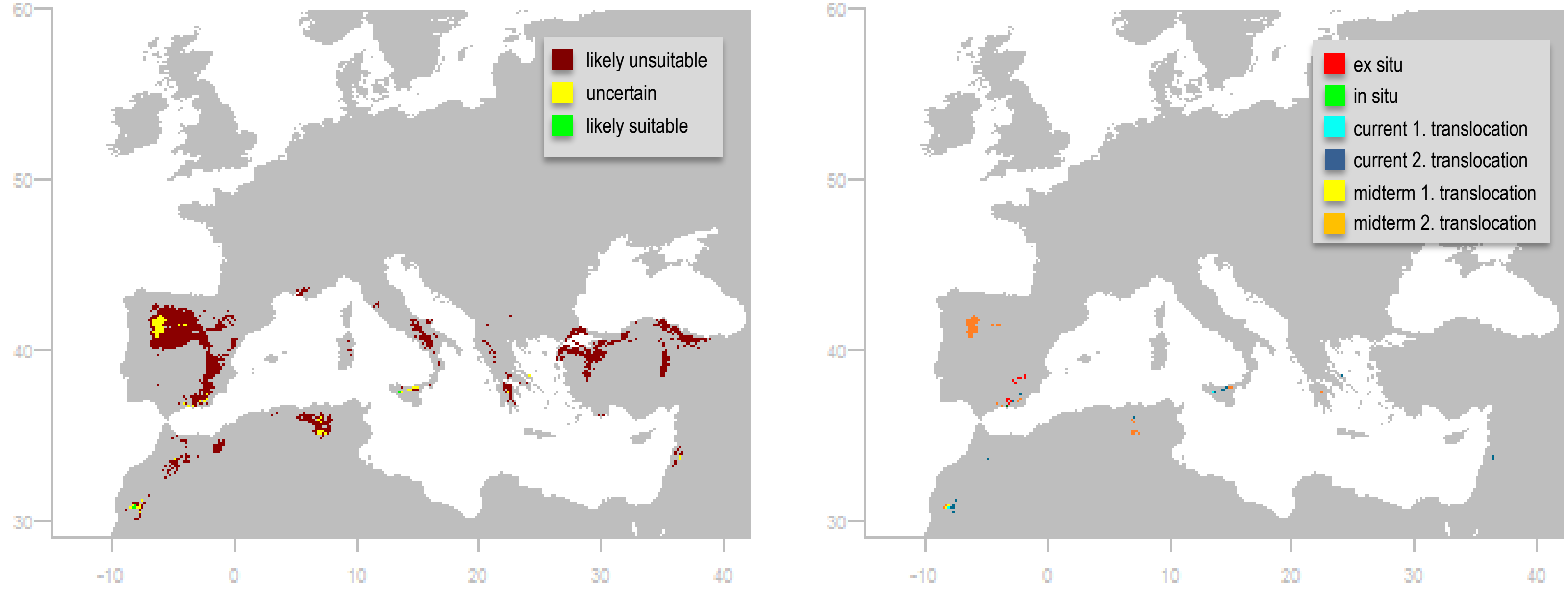
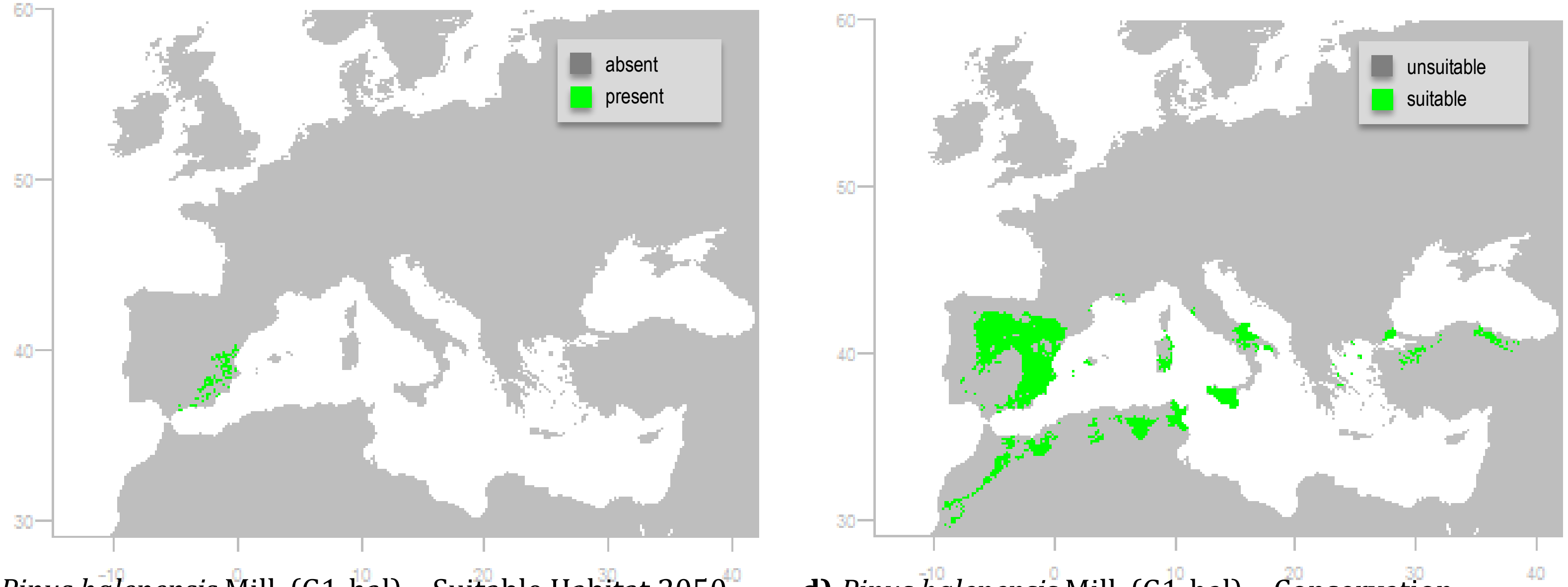


Figure S8:
a) Current distribution range of the Tunisia clade (G8-pin) of *Pinus pinaster* Ait.
b) Simulated habitat suitability under current climate for clade G8-pin.
c) Classified future (2050) habitat suitability map of G8-pin.
d) Conservation strategies proposed for clade G8-pin.

D: Figures for Pinus halepensis Mill. clades

a) *Pinus halepensis* Mill. (G1-hal) - Current Distribution

b) *Pinus halepensis* Mill. (G1-hal) – Suitable Habitat Current Climate



c) *Pinus halepensis* Mill. (G1-hal) – Suitable Habitat 2050

d) *Pinus halepensis* Mill. (G1-hal) – Conservation

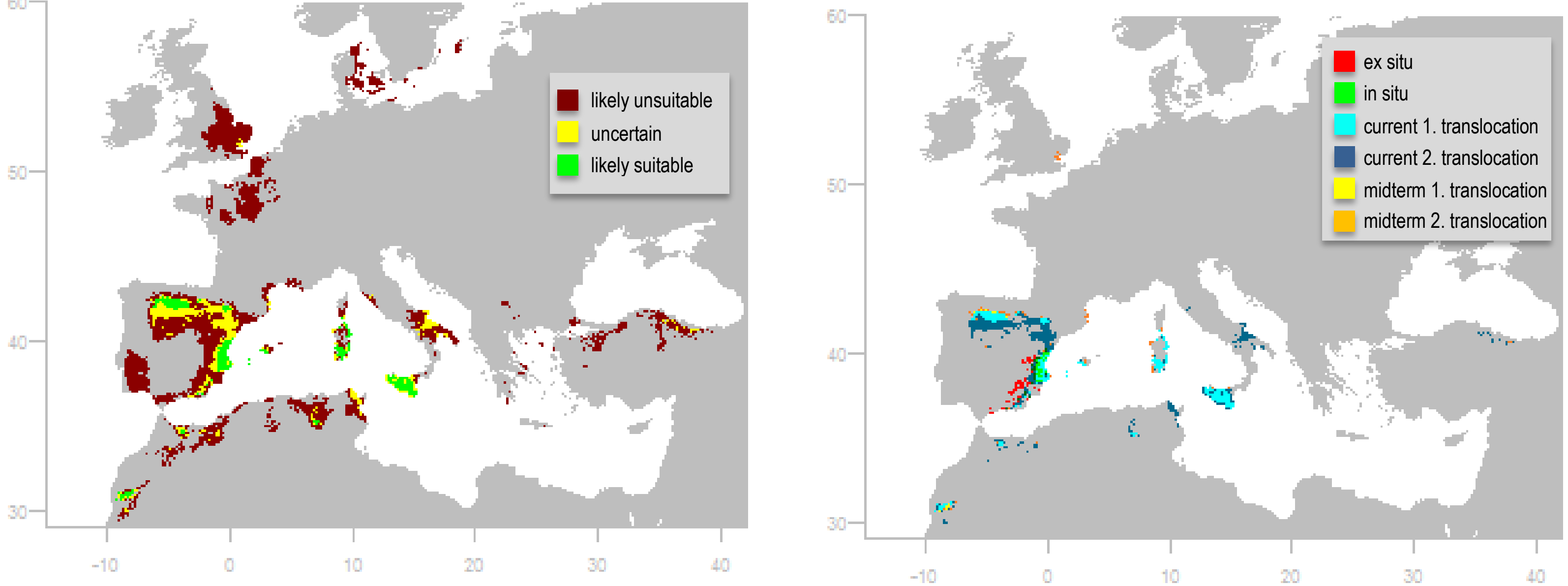


Figure S9:

a) Current distribution range of the Central and Southern Spain clade (G1-hal) of *Pinus halepensis* Mill.

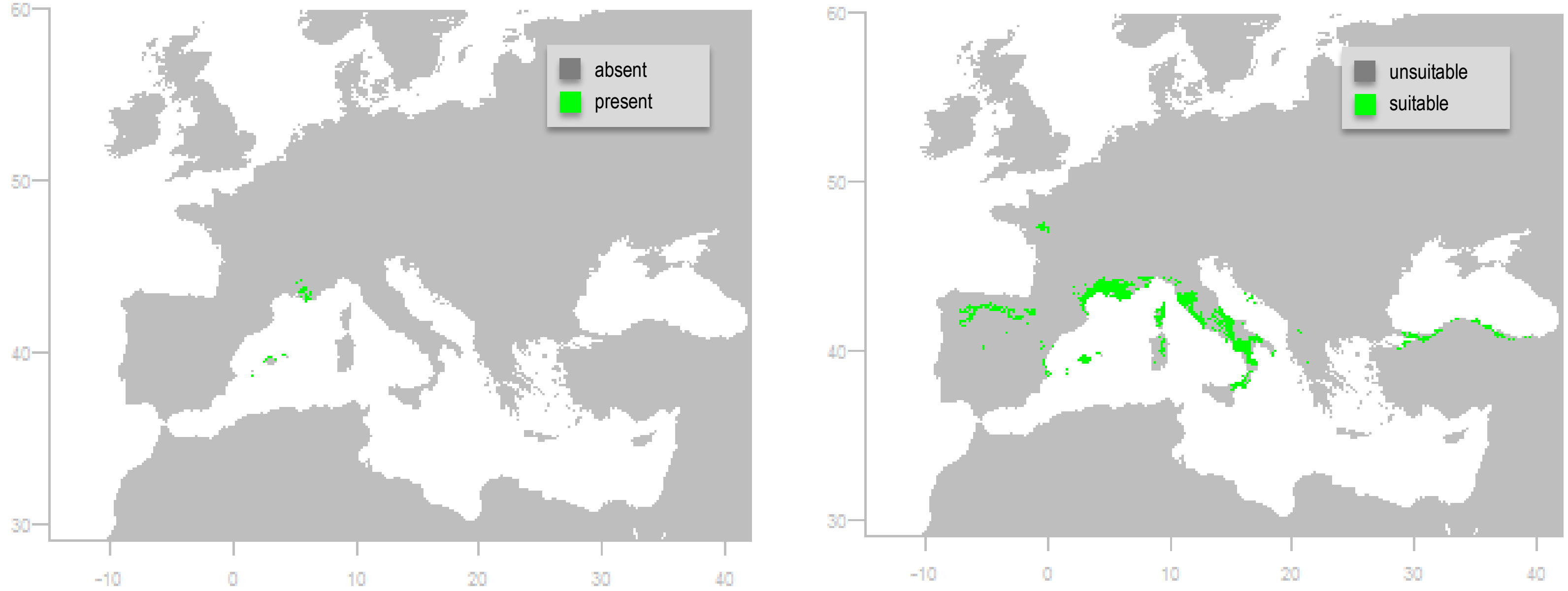
b) Simulated habitat suitability under current climate for clade G1-hal.

c): Classified future (2050) habitat suitability map of the Central and Southern Spain clade G1-hal.

d): Conservation strategies proposed for clade G1-hal.

a) *Pinus halepensis* Mill. (G2-hal) - Current Distribution

b) *Pinus halepensis* Mill. (G2-hal) – Suitable Habitat Current



c) *Pinus halepensis* Mill. (G2-hal) – Suitable Habitat 2050

d) *Pinus halepensis* Mill. (G2-hal) – Conservation

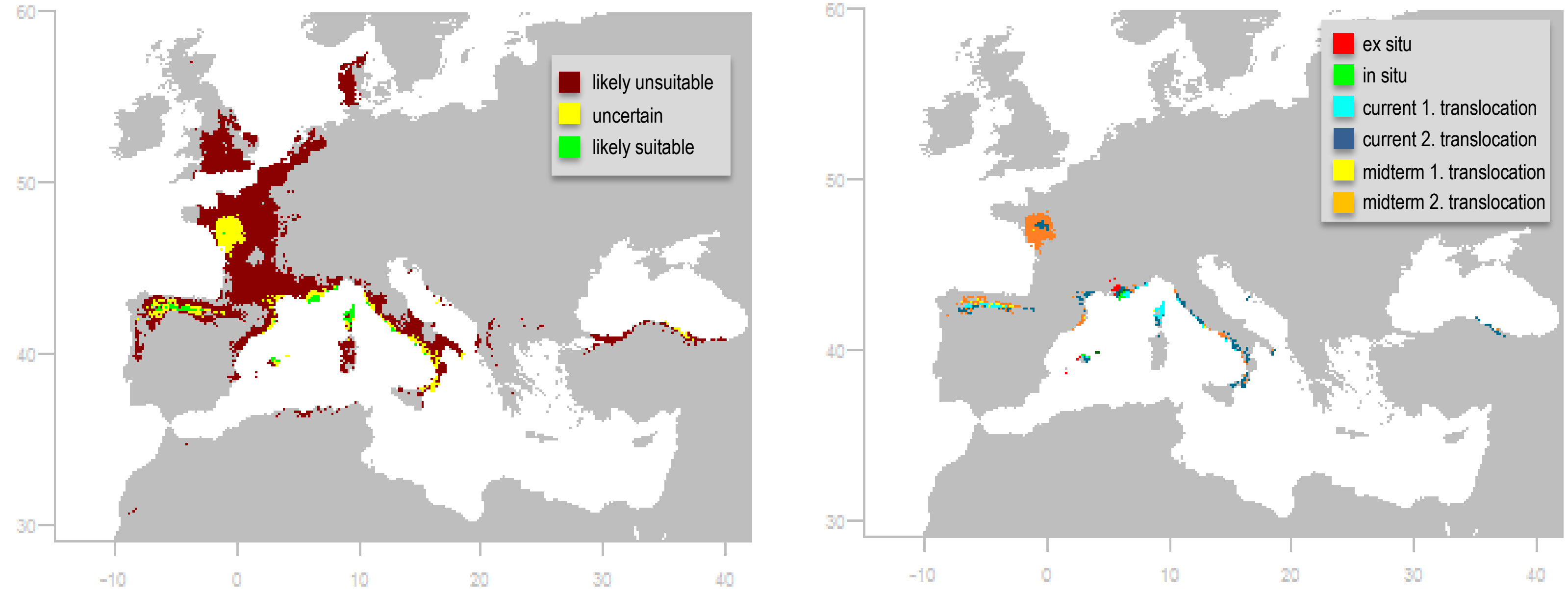


Figure S10 :

a) Current distribution range of the Balearian and Southern France clade (G2-hal) of *Pinus halepensis* Mill.

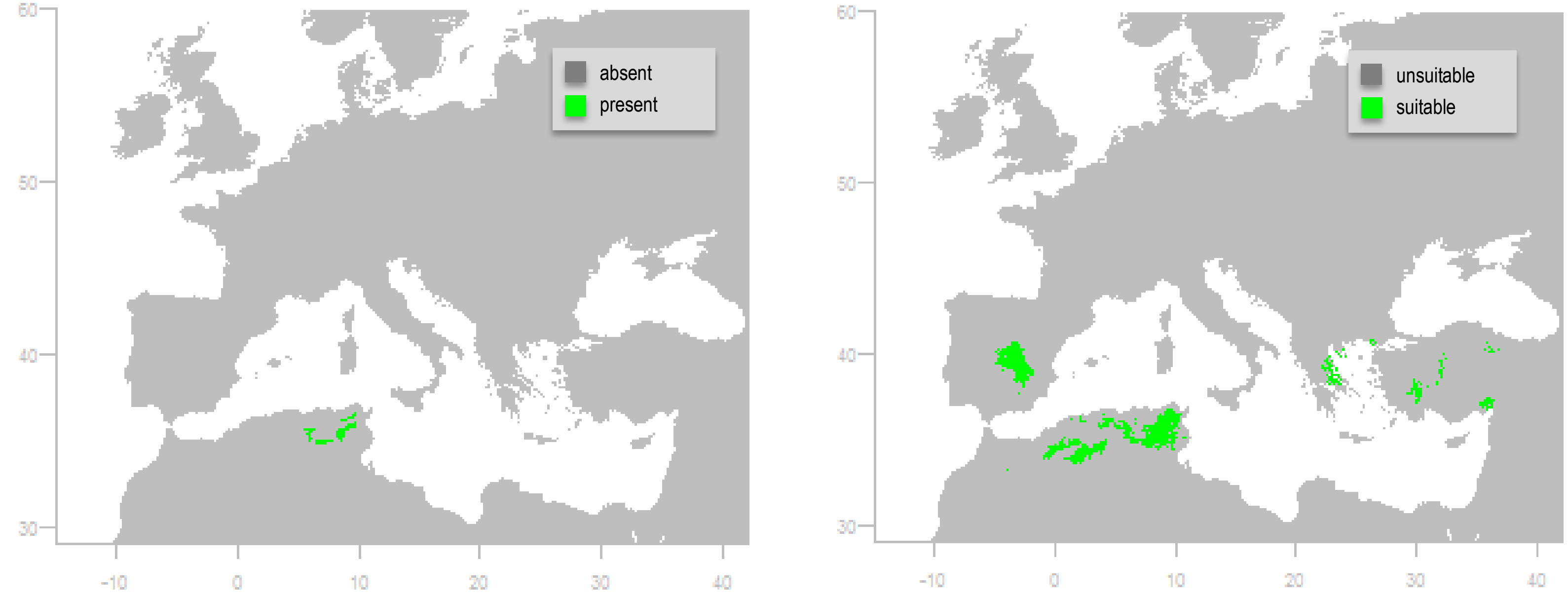
b) Simulated habitat suitability under current climate for clade G2-hal.

c): Classified future (2050) asuitability map of the Central and Southern Spain clade G2-hal.

d): Conservation strategies proposed for clade G2-hal.

a) *Pinus halepensis* Mill. (G3-hal) - Current Distribution

b) *Pinus halepensis* Mill. (G3-hal) – Suitable Habitat Current Climate



c) *Pinus halepensis* Mill. (G3-hal) – Suitable Habitat 2050

d) *Pinus halepensis* Mill. (G3-hal) – Conservation

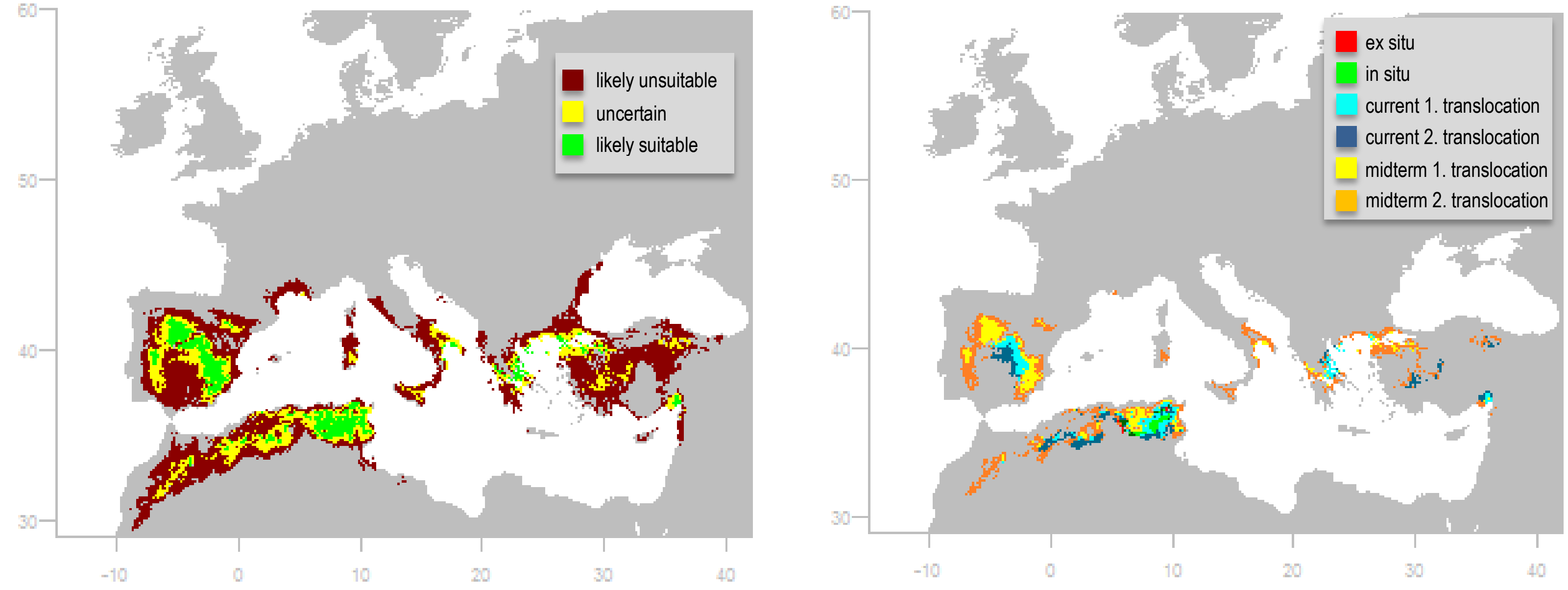


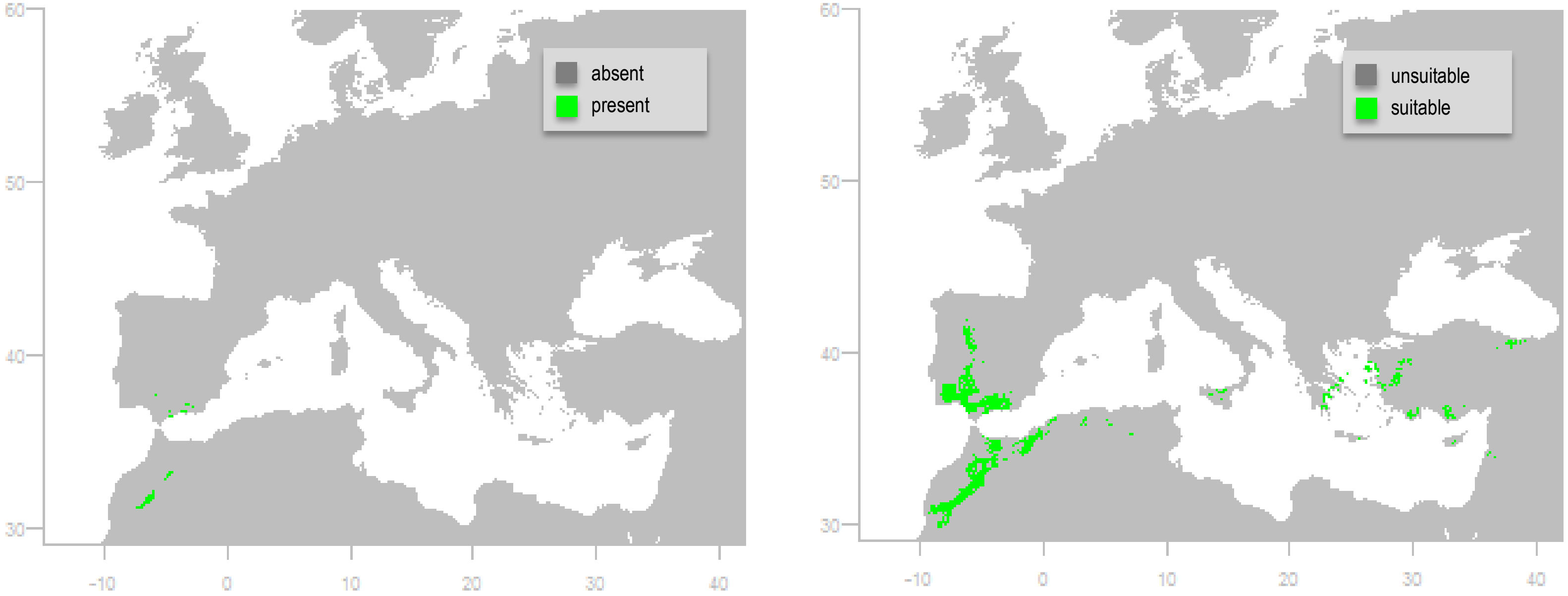
Figure S11 :

a) Current distribution range of Tunisian and Northern Italian clade (G3-hal) of *Pinus halepensis* Mill.

b) Simulated habitat suitability under current climate for clade G3-hal.

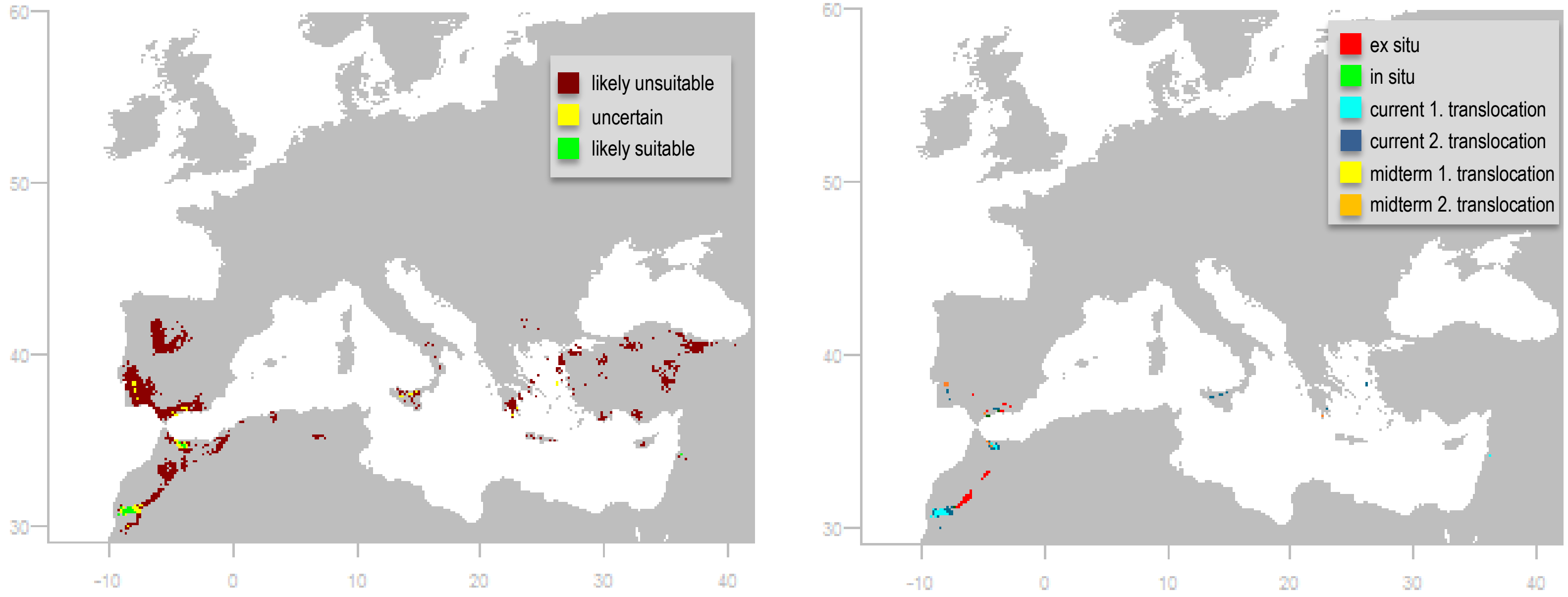
c): Classified future (2050) asuitability map of the Central and Southern Spain clade G3-hal.

d): Conservation strategies proposed for clade G3-hal.



a) *Pinus halepensis* Mill. (G4-hal) - Current Distribution

b) *Pinus halepensis* Mill. (G4-hal) – Suitable Habitat Current Climate



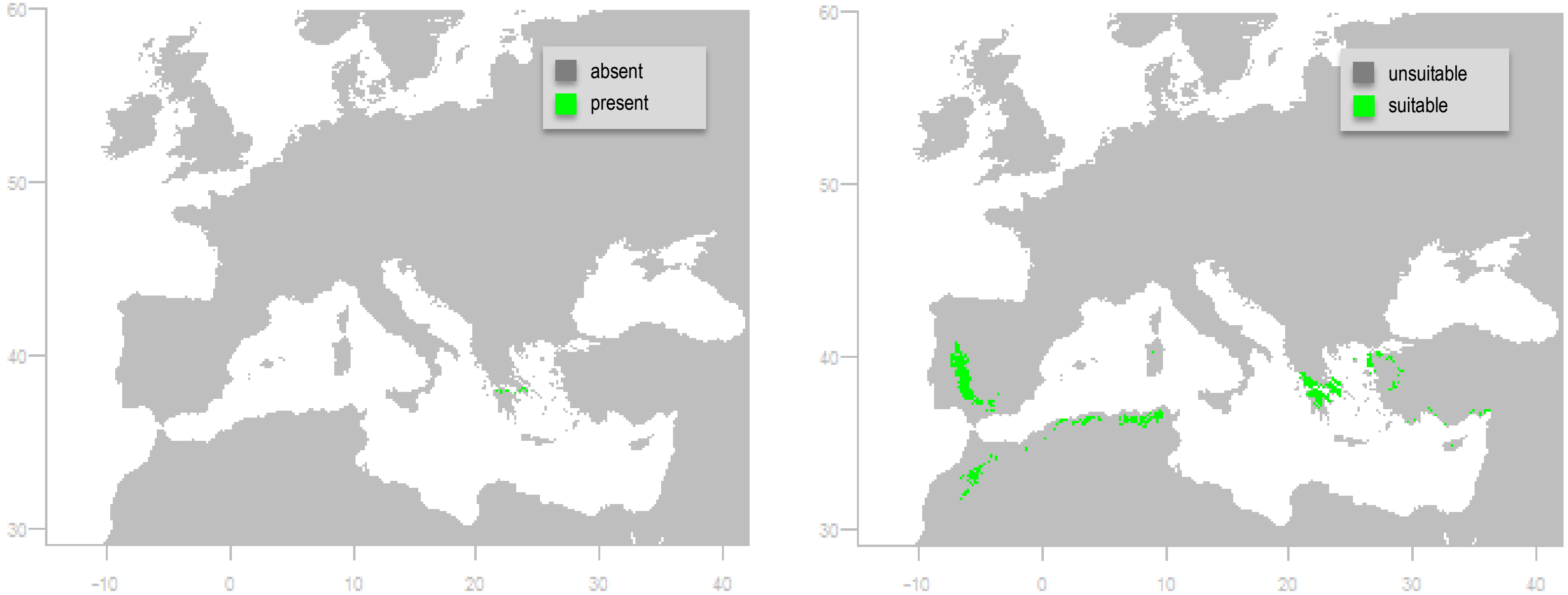
c) *Pinus halepensis* Mill. (G4-hal) – Suitable Habitat 2050

d) *Pinus halepensis* Mill. (G4-hal) – Conservation

Figure S12 : **a)** Current distribution range of Moroccan and Southern Spain clade (G4-hal) of *Pinus halepensis* Mill. **b)** Simulated habitat suitability under current climate for clade G4-hal. **c)**: Classified future (2050) suitability map of the Central and Southern Spain clade G4-hal. **d)**: Conservation strategies proposed for clade G4-hal.

a) *Pinus halepensis* Mill. (G5-hal) - Current Distribution

b) *Pinus halepensis* Mill. (G5-hal) – Suitable Habitat Current Climate



c) *Pinus halepensis* Mill. (G5-hal) – Suitable Habitat 2050

d) *Pinus halepensis* Mill. (G5-hal) – Conservation

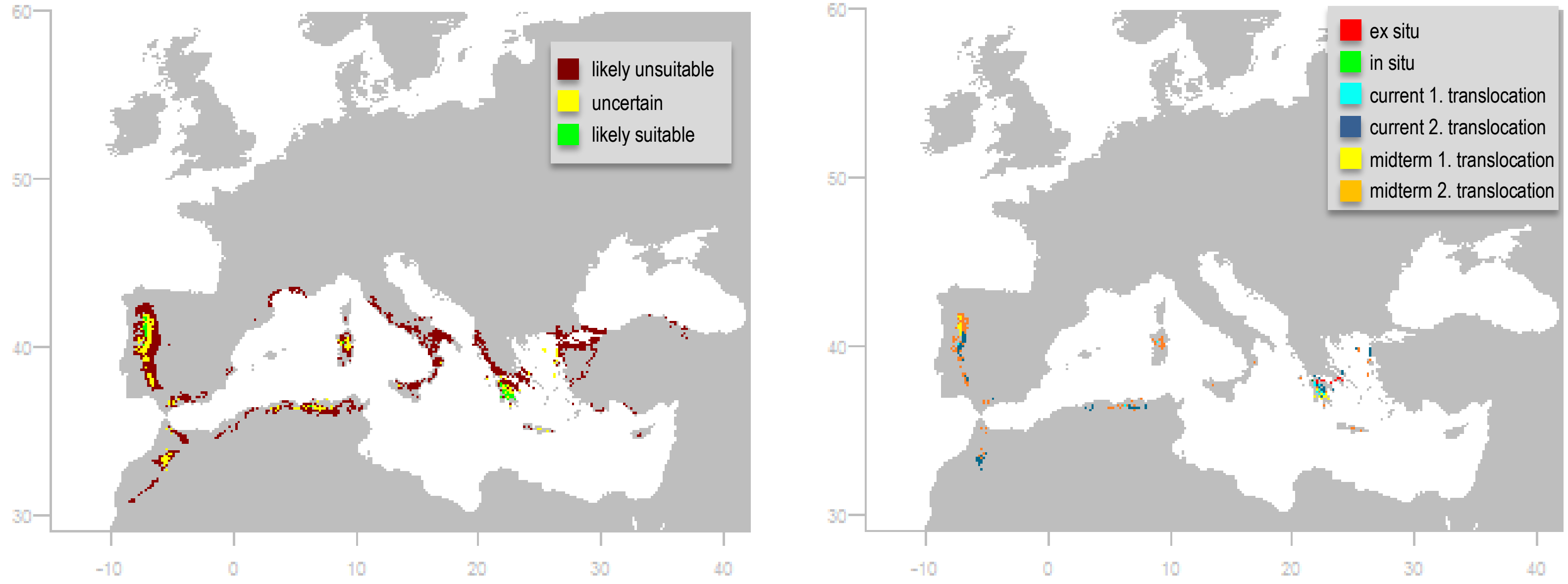
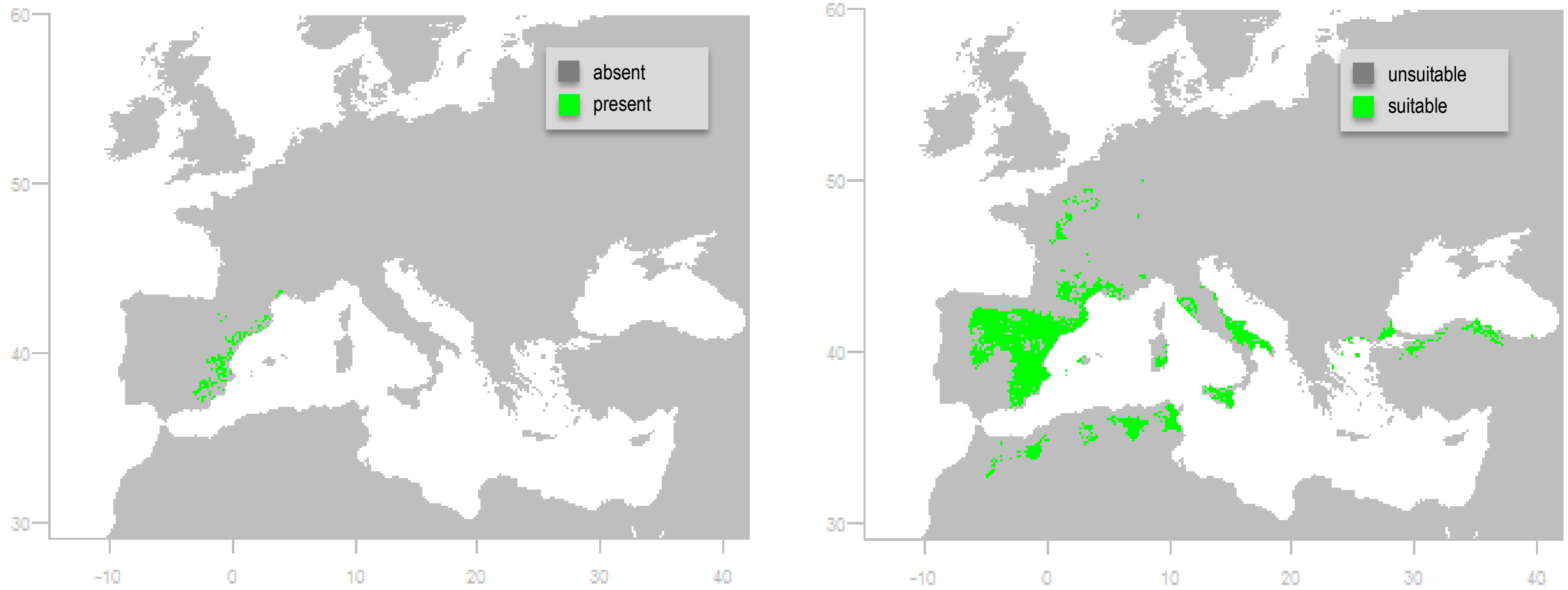


Figure S13 : **a)** Current distribution range Greek clade (G5-hal) of *Pinus halepensis* Mill. **b)** Simulated habitat suitability under current climate for clade G5-hal. **c)**: Classified future (2050) suitability map of the Central and Southern Spain clade G5-hal. **d)**: Conservation strategies proposed for clade G5-hal.

a) *Pinus halepensis* Mill. (G6-hal) - Current Distribution

b) *Pinus halepensis* Mill. (G6-hal) – Suitable Habitat Current



c) *Pinus halepensis* Mill. (G6-hal) – Suitable Habitat 2050

d) *Pinus halepensis* Mill. (G6-hal) – Conservation

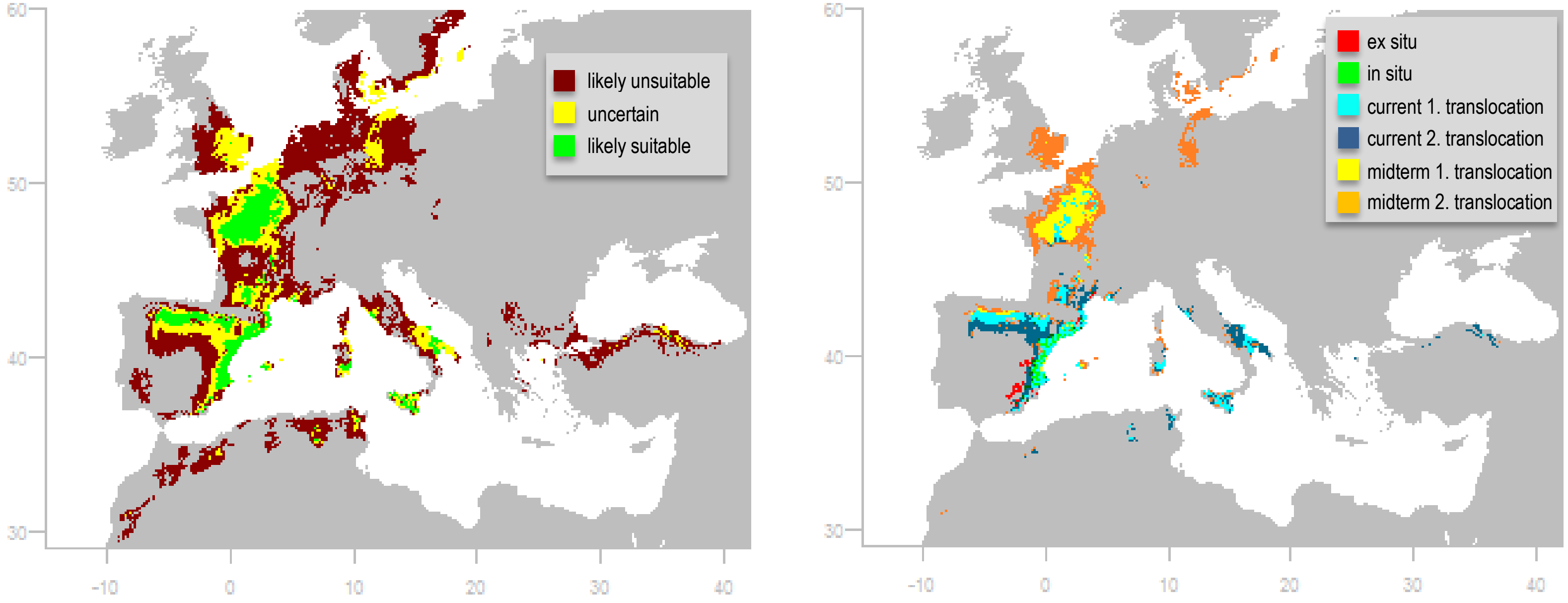


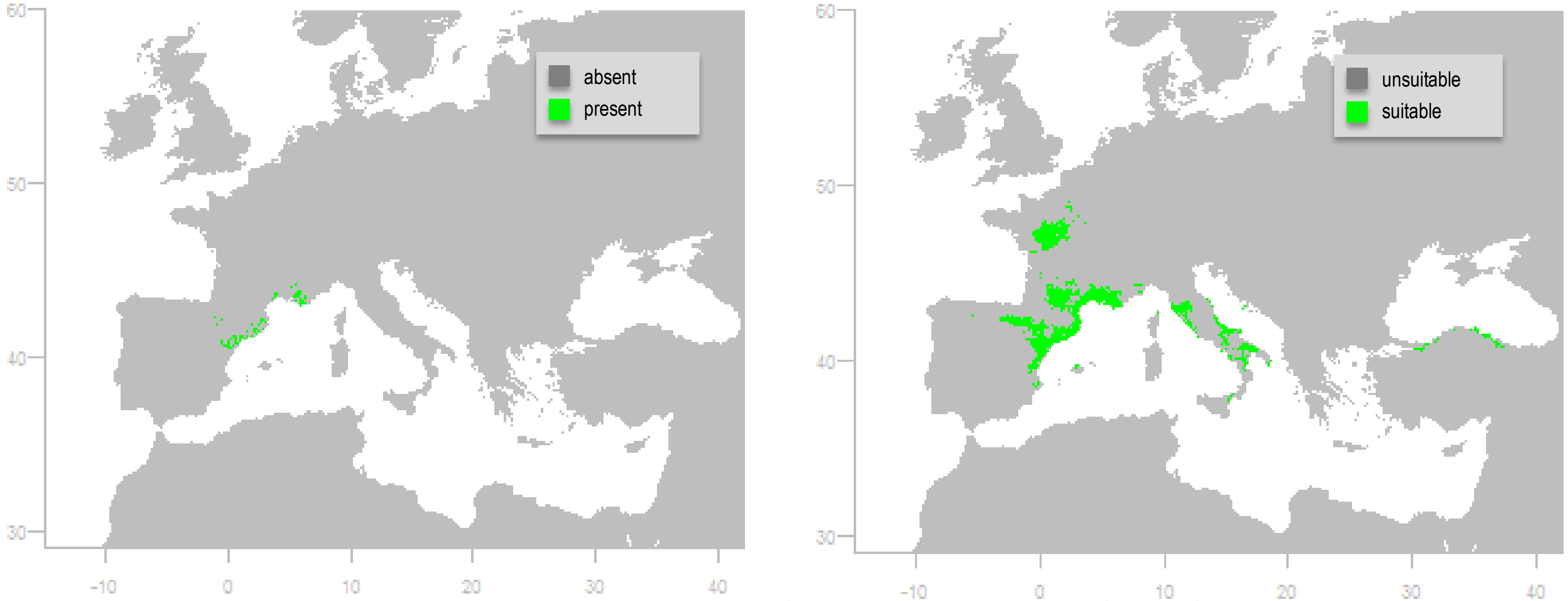
Figure S14 :

a) Current distribution range Central-Northern Spain clade (G6-hal) of *Pinus halepensis* Mill.

b) Simulated habitat suitability under current climate for clade G6-hal.

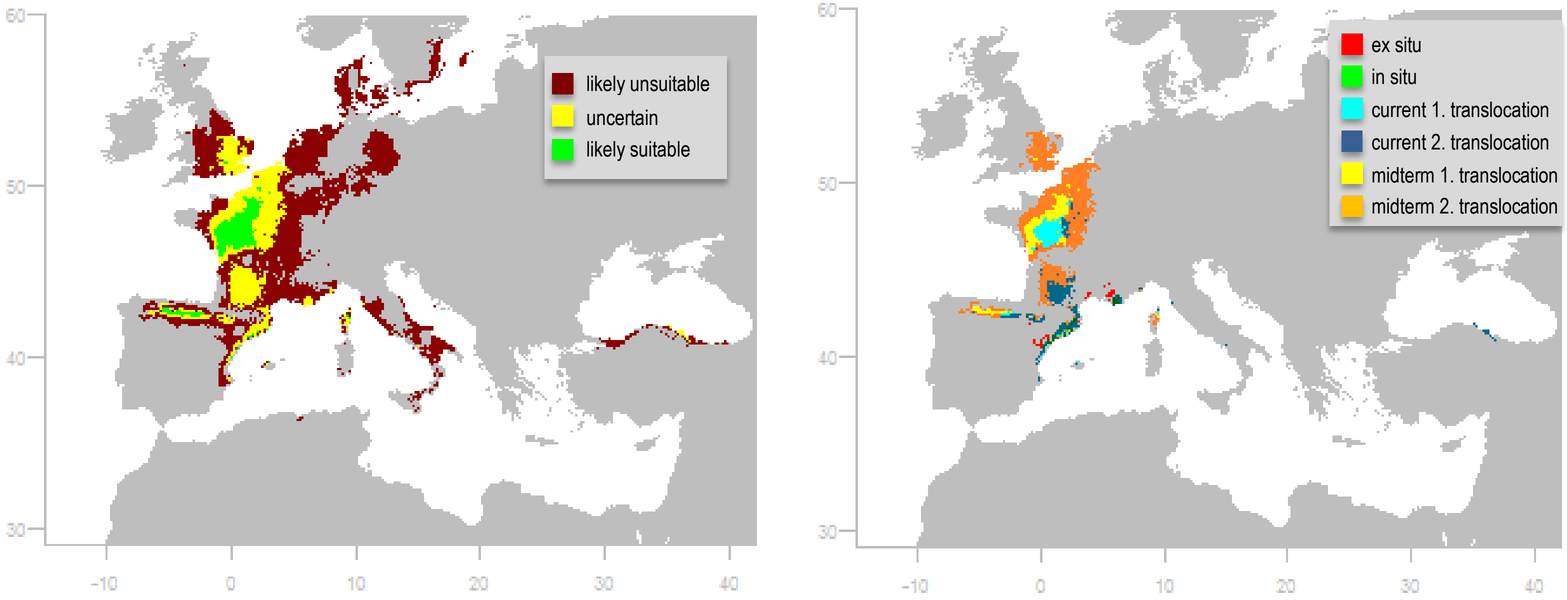
c): Classified future (2050) suitability map of the Central and Southern Spain clade G6-hal.

d): Conservation strategies proposed for clade G6-hal.



a) *Pinus halepensis* Mill. (G7-hal) - Current Distribution

b) *Pinus halepensis* Mill. (G7-hal) – Suitable Habitat Current Climate



c) *Pinus halepensis* Mill. (G7-hal) – Suitable Habitat 2050

d) *Pinus halepensis* Mill. (G7-hal) – Conservation

Figure S15 :

a) Current distribution range Northern Spain-Southern France clade (G7-hal) of *Pinus halepensis* Mill.

b) Simulated habitat suitability under current climate for clade G7-hal.

c): Classified future (2050) suitability map of the Central and Southern Spain clade G7-hal.

d): Conservation strategies proposed for clade G7-hal.

E: Summary table of areas assigned to different conservation strategies

Table S6: Areas (Km²) corresponding to the territories proposed for the different conservation strategies within each clade for both *Pinus pinaster* Ait. and *Pinus halepensis* Mill.

	In situ	***In situ*** **with monitoring**	***Ex situ***	**%** ***Ex situ***	**Current 1 translocation**	**Current 2 translocation**	**Midterm 1 translocation**	**Midterm 2 translocation**
G1-pin	17705.16	8204.38	12551.22	32.63	19900.94	11155.76	495.22	7012.78
G2-pin	4018.92	2509.93	1087.43	14.28	56344.12	91077.41	12576.06	80793.09
G3-pin	0.00	5546.33	5461.39	49.61	152.78	5801.86	2436.11	12412.18
G4-pin	25.27	1033.07	2474.07	70.04	2942.59	16597.06	17.62	5083.37
G5-pin	0.00	234.51	5551.36	95.95	1126.91	17304.47	2377.91	15864.50
G6-pin	0.00	0.00	8442.64	100.00	183.11	532.33	39.89	303.46
G7-pin	0.00	74.32	3245.49	97.76	624.98	4616.07	252.42	13796.36
G8-pin	914.61	1797.52	81.90	2.93	1661.48	8185.67	3286.68	19431.03
G1-hal	6083.29	4973.49	11938.10	51.92	51656.41	93759.80	3462.11	15869.45
G2-hal	1428.37	1081.02	2573.68	50.63	14312.61	32168.77	3254.16	48487.86
G3-hal	11798.66	3653.03	1845.59	10.67	59856.54	68721.48	87104.06	230536.26
G4-hal	0.00	622.22	8174.68	92.93	5044.45	10857.58	95.36	1644.97
G5-hal	80.76	500.10	2388.70	80.44	2883.85	15812.66	4617.21	24701.42
G6-hal	16002.16	5937.70	10021.55	31.36	93155.30	113639.69	75152.99	207605.12
G7-hal	953.86	5438.08	8118.76	55.95	27357.19	50344.32	39672.77	179632.45

APPENDIX S4:

Dynamic Conservation Units: Classification within clades and Exposure assessment

Table S7: *Pinus pinaster* Ait. Exposure assessment of its dynamic conservation units (DCUs)

DCU Code	Clade	Future Suitability 2050
ESP00018	Southern Spanish group - G7	Likely unsuitable
ESP00016	Southern Spanish group - G7	Likely unsuitable
ESP00007	Southern Spanish group - G7	Likely unsuitable
ESP00017	Southern Spanish group - G7	Likely unsuitable
ESP00042	Southern Spanish group - G7	Uncertain
ESP00043	Southern Spanish group - G7	Uncertain
ESP00025	Not within Euforgen limits (G7)	Uncertain
ESP00045	Not within Euforgen limits (G7)	Likely unsuitable
ESP00035	Southern Spanish group - G7	Likely unsuitable
ESP00038	Southern Spanish group - G7	Likely unsuitable
ESP00039	Eastern Spanish group - G5	Likely suitable
ESP00037	Central Spanish group - G6	Likely unsuitable
ESP00032	Eastern Spanish group - G5	Uncertain
ESP00044	Eastern Spanish group - G5	Uncertain
ESP00030	Not classified	Not classified
ESP00023	Central Spanish group - G6	Likely unsuitable
ESP00033	Not within Euforgen limits (G5)	Uncertain
ESP00034	Not within Euforgen limits (G5)	Uncertain
ESP00006	Not within Euforgen limits (G5)	Uncertain
ESP00031	Eastern Spanish group - G5	Likely unsuitable
ESP00024	Not classified	Not classified
ESP00020	Not classified	Not classified
ESP00022	Not classified	Not classified
ESP00029	Not classified	Not classified
ITA00135	Eastern group - G2	Likely unsuitable
ITA00019	Eastern group - G2	Likely unsuitable
ITA00172	Eastern group - G2	Likely unsuitable
ITA00073	Eastern group - G2	Uncertain
ITA00242	Eastern group - G2	Uncertain
ITA00136	Eastern group - G2	Uncertain
ITA00152	Not within Euforgen limits (G2)	Uncertain
ITA00011	Eastern group - G2	Likely suitable
ITA00104	Eastern group - G2	Likely suitable
ITA00132	Eastern group - G2	Likely suitable
ITA00097	Eastern group - G2	Likely suitable
ITA00060	Eastern group - G2	Likely suitable
FRA00048	Atlantic France group - G3	Uncertain

FRA00049	Atlantic France group - G3	Uncertain
ITA00203	Eastern group - G2	Likely suitable
ITA00059	Not within Euforgen limits (G2)	Coordinates outside study area
FRA00050	Atlantic France group - G3	Uncertain
FRA00051	Atlantic France group - G3	Uncertain

Table S8: *Pinus halepensis* Mill Exposure assessment of its dynamic conservation units (DCUs)

DCU Code	Clade	Future Suitability 2050 (Clade 1)	Future Suitability 2050 (Clade 2)
GRC00001	Greek clade (G5)	Likely unsuitable	-
GRC00002	Not within Euforgen limits (G3 & G5)	Coordinates outside study area	Coordinates outside study area
GRC00003	Not within Euforgen limits (G3 & G5)	Coordinates outside study area	Coordinates outside study area
ITA00027	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Coordinates outside study area	Coordinates outside study area
ITA00107	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Likely suitable	Likely unsuitable
ITA00108	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Likely suitable	Likely unsuitable
ITA00109	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Likely suitable	Likely unsuitable
ITA00046	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Coordinates outside study area	Coordinates outside study area
ITA00110	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Likely suitable	Likely unsuitable
ITA00111	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Likely suitable	Likely unsuitable
ITA00112	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Likely suitable	Likely unsuitable
ITA00113	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Coordinates outside study area	Coordinates outside study area
ITA00114	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Coordinates outside study area	Coordinates outside study area
ITA00025	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Coordinates outside study area	Coordinates outside study area
ITA00115	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Likely unsuitable	Likely unsuitable
ITA00133	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Likely unsuitable	Likely unsuitable
ESP00003	Central and northern Spain clade (G6) & Northern Spain-southern France clade (G7)	Uncertain	Likely unsuitable
ITA00076	Not within Euforgen limits (G3 & G5)	Likely unsuitable	Likely unsuitable
ITA00072	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Likely unsuitable	Likely unsuitable
ITA00075	Tunisian and northern Italian clade	Likely unsuitable	Likely unsuitable

	(G3) & Greek clade (G5)		
BIH00065	Tunisian and Northern Italian clade (G3) & Greek clade (G5)	Likely unsuitable	Likely unsuitable
ITA00007	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Likely unsuitable	Likely unsuitable
ITA00242	Tunisian and northern Italian clade (G3) & Greek clade (G5)	Likely unsuitable	Likely unsuitable
ITA00061	Balearic and southern France clade (G2) & Northern Spain and southern France clade (G7)	Likely suitable	Likely unsuitable
ITA00179	Balearic and southern France clade (G2) & Northern Spain and southern France clade (G7)	Likely suitable	Likely unsuitable
ITA00165	Balearic and southern France clade (G2) & Northern Spain and southern France clade (G7)	Likely unsuitable	Likely unsuitable

Table S9: *Pinus pinaster* Ait. Exposure assessment of its genetically-homogeneous conservation relevant population groups

Population name	Code	Genetic clade	Conservation groups	Future Suitability 2050
Pleucadec	PLE	Atlantic France group - G3	FrAtl_1	Likely unsuitable
Hourtin	HOU	Atlantic France group - G3	FrAtl_2	Uncertain
Mimizan	MIM	Atlantic France group - G3	FrAtl_1	Uncertain
Olonne sur Mer	OLO	Atlantic France group - G3	FrAtl_2	Coordinates outside study area
Petrocq	PET	Atlantic France group - G3	FrAtl_2	Uncertain
St-Jeans des Monts	STJ	Atlantic France group - G3	FrAtl_2	Coordinates outside study area
Leverdon	VER	Atlantic France group - G3	FrAtl_2	Likely unsuitable
Alto Llama	ALT	Atlantic Iberian Peninsula group - G1	SpAtl	Likely unsuitable
Arenas S.Pedro	ARM	Atlantic Iberian Peninsula group - G1	SpAtl	Likely unsuitable
Cadavedo	CAD	Atlantic Iberian Peninsula group - G1	SpAtl	Likely unsuitable
Castropol	CAS	Atlantic Iberian Peninsula group - G1	SpAtl	Likely unsuitable
Lamu±o	LAM	Atlantic Iberian Peninsula group - G1	SpAtl	Likely unsuitable
Leiria	LEI	Atlantic Iberian Peninsula group - G1	SpAtl	Likely suitable
Pto. Vega	PUE	Atlantic Iberian Peninsula group - G1	SpAtl	Uncertain
S.Cipriano	SAC	Atlantic Iberian Peninsula group - G1	SpAtl	Likely suitable
Sergurde	SEG	Atlantic Iberian Peninsula group - G1	SpAtl	Likely suitable
Sier. Barcia	SIE	Atlantic Iberian Peninsula group - G1	SpAtl	Likely unsuitable
Tamrabta	TAM	Moroccan group - G4	Mor	Likely unsuitable
Boniches	BON	Eastern Spanish group - G5	CSp_2	Likely unsuitable
Olba	OLB	Eastern Spanish group - G5	CSp_2	Uncertain
S.Leonardo	SAL	Central Spanish group - G6	CSp_1	Likely unsuitable
Valdemaqueda	VAL	Central Spanish group - G6	CSp_2	Likely unsuitable
Pineta	PIE	Eastern group - G2	Pie	Uncertain
Bayubas	BAY	Central Spanish group - G6	CSp_1	Likely unsuitable
Coca	COC	Central Spanish group - G6	CSp_1	Likely unsuitable
Armayán	ARN	Central Spanish group - G6	CSp_2	Likely unsuitable
Cuellar	CUE	Central Spanish group - G6	CSp_1	Likely unsuitable
Oria	ORI	Southern Spanish group- G7	Ori	Uncertain
Cenicientos	CEN	Central Spanish group - G6	CSp_2	Likely unsuitable
Pinia	PIA	Eastern group - G2	Pia	Likely unsuitable
Quatretonda	QUA	Southern Spanish group - G7	Qua	Likely unsuitable

Supporting information Chapter 3:

APPENDIX S1:

Details on meteorological stations and climatic data

APPENDIX S1:

Details on meteorological stations and climatic data

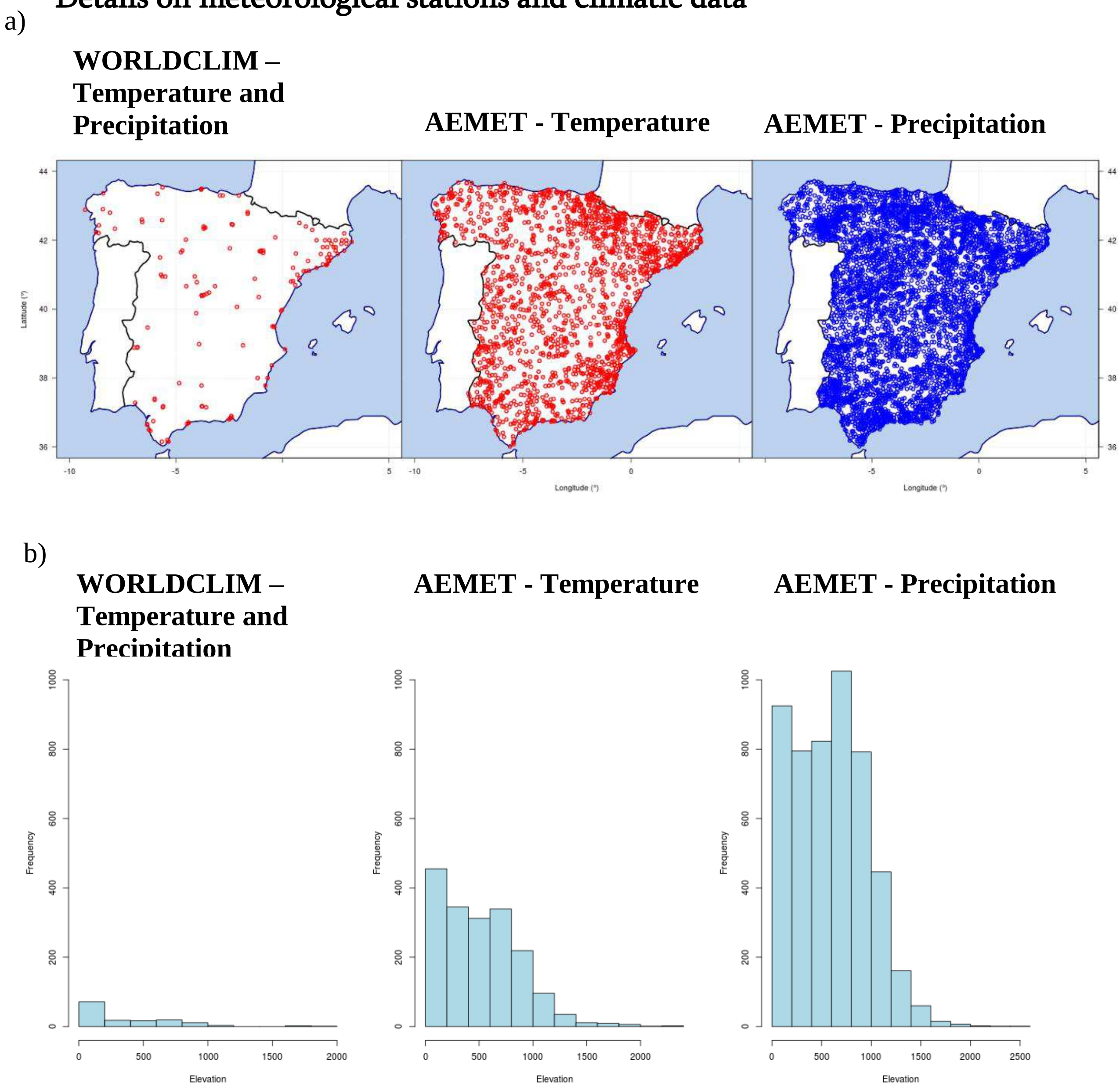


Fig S1: Geographical representation (a) and histogram of elevations (b) of the 142 meteorological stations employed for climatic interpolations in WORLDCLIM database (temperature and precipitation; Hijmans *et al.,* 2005) and the 1830 (temperature) and 5053 (precipitation) meteorological stations provided by the Spanish Meteorological Agency (AEMET) across the Spanish Iberian Peninsula.

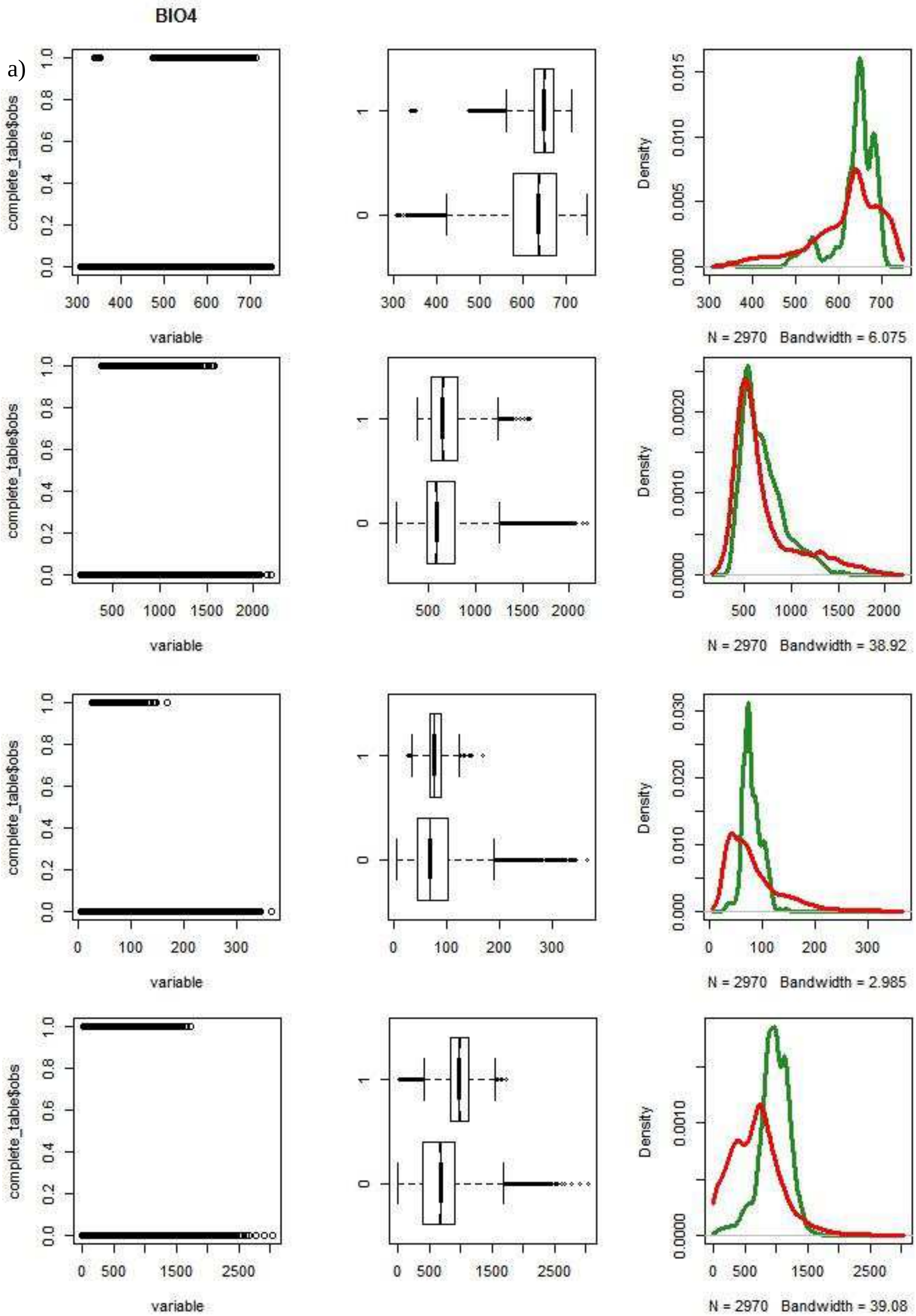
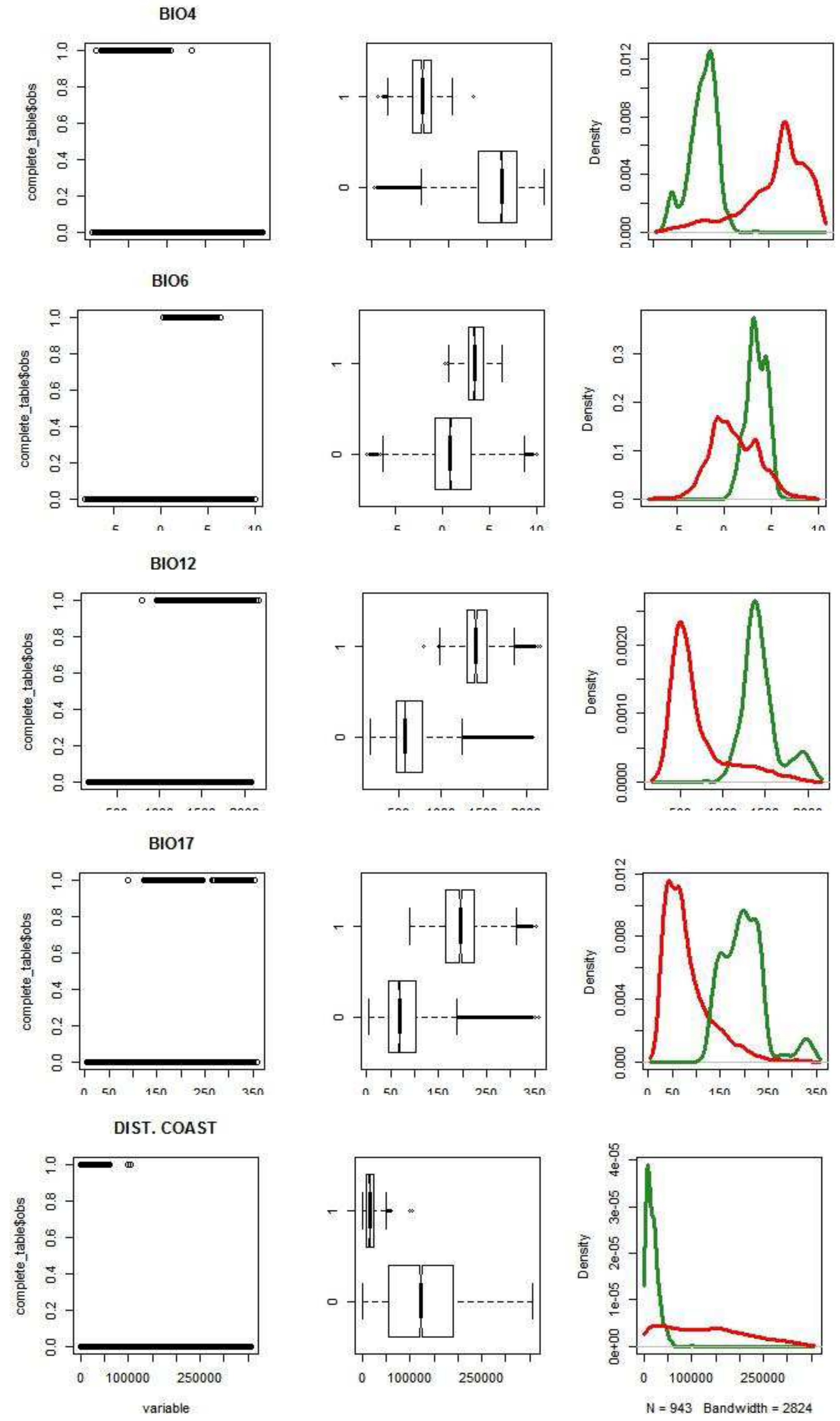


Fig S2: Selected environmenal variables for *Pinus pinaster* Ait. (a) and pitch canker disease (b) to be included within their species distribution models. 1/0 values as well as green and red lines correspond to presence and pseudoabsence records respectively.

b)


BIO4
complete_table$obs
0.0
0.2
0.4
0.6
0.8
1.0
1
0
Density
0.000
0.004
0.008
0.012
BIO6
complete_table$obs
0.0
0.2
0.4
0.6
0.8
1.0
1
0
Density
0.0
0.1
0.2
0.3
BIO12
complete_table$obs
0.0
0.2
0.4
0.6
0.8
1.0
1
0
Density
0.0000
0.0010
0.0020
BIO17
complete_table$obs
0.0
0.2
0.4
0.6
0.8
1.0
0
50
150
250
350
1
0
Density
0.000
0.004
0.008
0.012
DIST. COAST
complete_table$obs
0.0
0.2
0.4
0.6
0.8
1.0
0
100000
250000
variable
1
0
Density
0e+00
1e-05
2e-05
3e-05
4e-05
N = 943 Bandwidth = 2824

Supporting information Chapter 4:

APPENDIX S1:

Environmental variable selection, Environmental marginality figures and Geographical and environmental indices for genetically characterized populations

APPENDIX S2:

Geographic and environmental indices for genetically characterized populations

APPENDIX S1

Environmental variable selection, Environmental marginality figures and Geographical and environmental indices for genetically characterized populations

Variable selection for Species Distribution Models

We downloaded the 19 bioclimatic variables available in WORLDCLIM (Hijmans *et al.*, 2005) representative of the period 1950-2000 for the analysis.

In order to avoid multicollinearity effects, we retained variables with Pearson correlations lower than 0.75 as the use of simple methods based on rules of thumb have proved to be as effective as more complicated methods (Dormann *et al.*, 2013). This threshold value is a bit less restrictive than the common value of 0.70 but more flexible thresholds have also been used in literature (see Elith *et al.,* 2006 for an example). Among highly correlated variables we kept the one with the highest explained deviance scores (D^2) when individually fitted in a Generalized Linear Model (GLM; McCullagh & Nelder, 1989). We performed this analysis individually for each species and selected four or five variables, both temperature- and precipitation-related, to characterize their bioclimatic niche. For all the species, we automatically discarded BIO8 and BIO9 as the steep gradient shown by these variables, in which very often two adjacent cells are characterized by extremely different values within the study area for no obvious reason, may lead to artefacts in the SDM output maps. Finally we performed a Variance Inflation Factor (VIF) test and adjusted variable selection to ensure that all VIF values were below a threshold value of 10. The similar patterns detected in D^2 values in *P. pinaster* /*P. halepensis* and *P. picea* / *P. sylvestris* enabled the selection of the same set of variables for each pair of species while *P. nigra, P. pinea* and *A. alba* had an independent bioclimatic variable set. The final set of relevant weakly correlated variables for each species is shown in Table S1.

Table S1: Set of relevant and weakly correlated bioclimatic predictors selected for each target species

Species	**Bioclimatic predictors**				
Pinus halepensis Pinus pinaster	BIO1 (Annual Mean Temperature)	BIO4 (Temperature Seasonality)	BIO18 (Precipitation of Warmest Quarter)	BIO19 (Precipitation of Coldest Quarter)	-
Pinus nigra	BIO2 (Mean diurnal range)	BIO4 (Temperature Seasonality)	BIO11 (Mean Temperature of Coldest Quarter)	BIO19 (Precipitation of Coldest Quarter)	-
Pinus pinea	BIO1 (Annual Mean Temperature)	BIO4 (Temperature Seasonality)	BIO15 (Precipitation seasonality)	BIO18 (Precipitation of Warmest Quarter)	BIO19 (Precipitation od Coldest Quarter)
Abies alba	BIO3 (Isothermality)	BIO7 (Temperature Annual Range)	BIO10 (Mean Temperature of Warmest Quarter)	BIO16 (Precipitation of Wettest Quarter)	BIO18 (Precipitation of Warmest Quarter)
Pinus sylvestris Picea abies	BIO3 (Isothermality)	BIO4 (Temperature Seasonality)	BIO10 (Mean Temperature of Warmest Quarter)	BIO15 (Precipitation seasonality)	BIO18 (Precipitation of Warmest Quarter)

References

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., Mcclean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D. & Lautenbach, S. (2013) Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **36**, 027–046.

Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-pereira, R., Schapire, R.E., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species ' distributions from occurrence data. *Ecography*, **29**, 129–151.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

McCullagh, P. & Nelder, J.A. (1989) *Generalized linear models*, (ed. by C. and Hall).

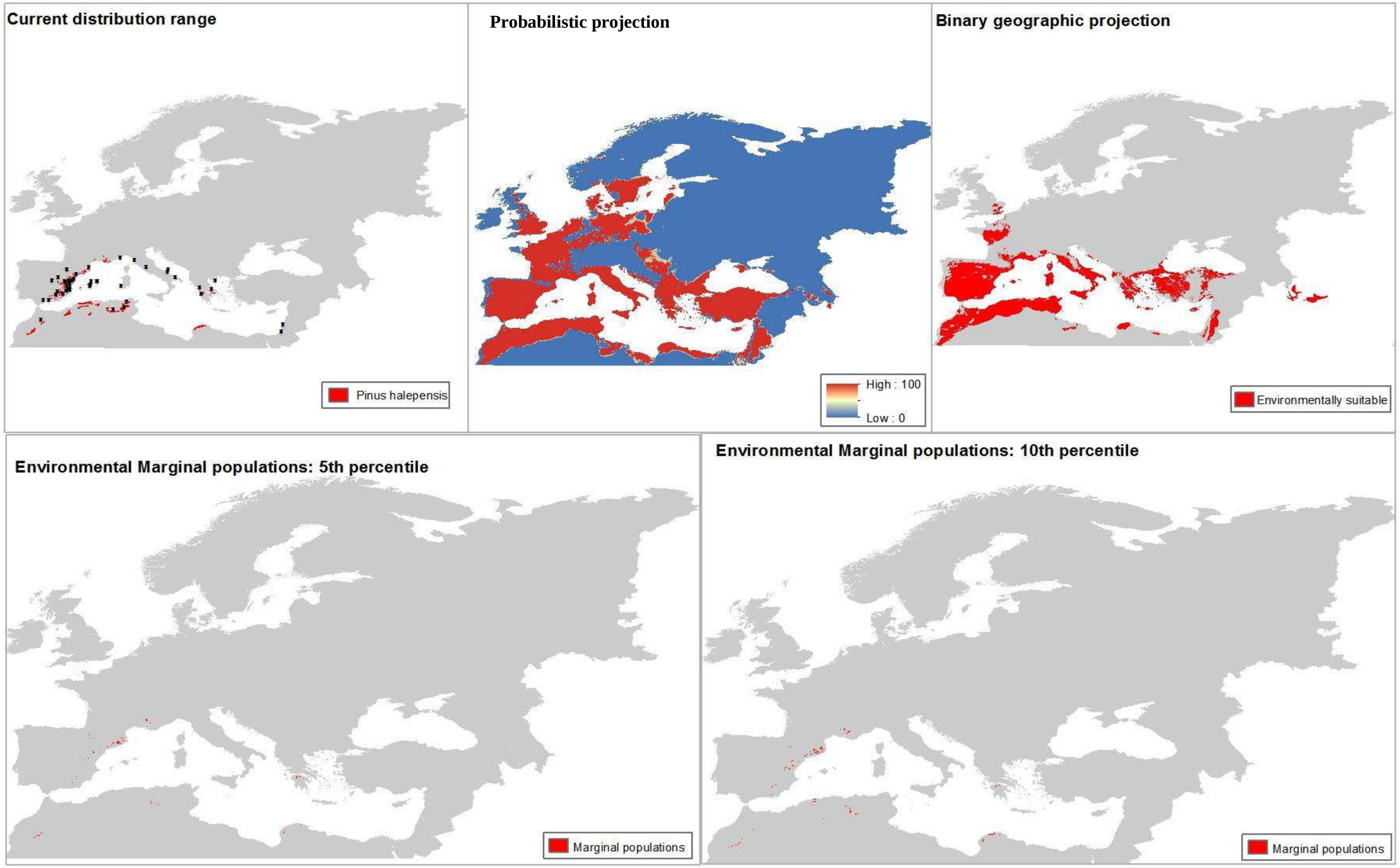


Fig S1: Current distribution, probabilistic and binary geographic projections from Species Distribution Models, and environmental marginal populations (two different thresholds: 5th and 10th percentiles) of *Pinus halepensis* Mill. Black dots correspond to genetically characterized populations.

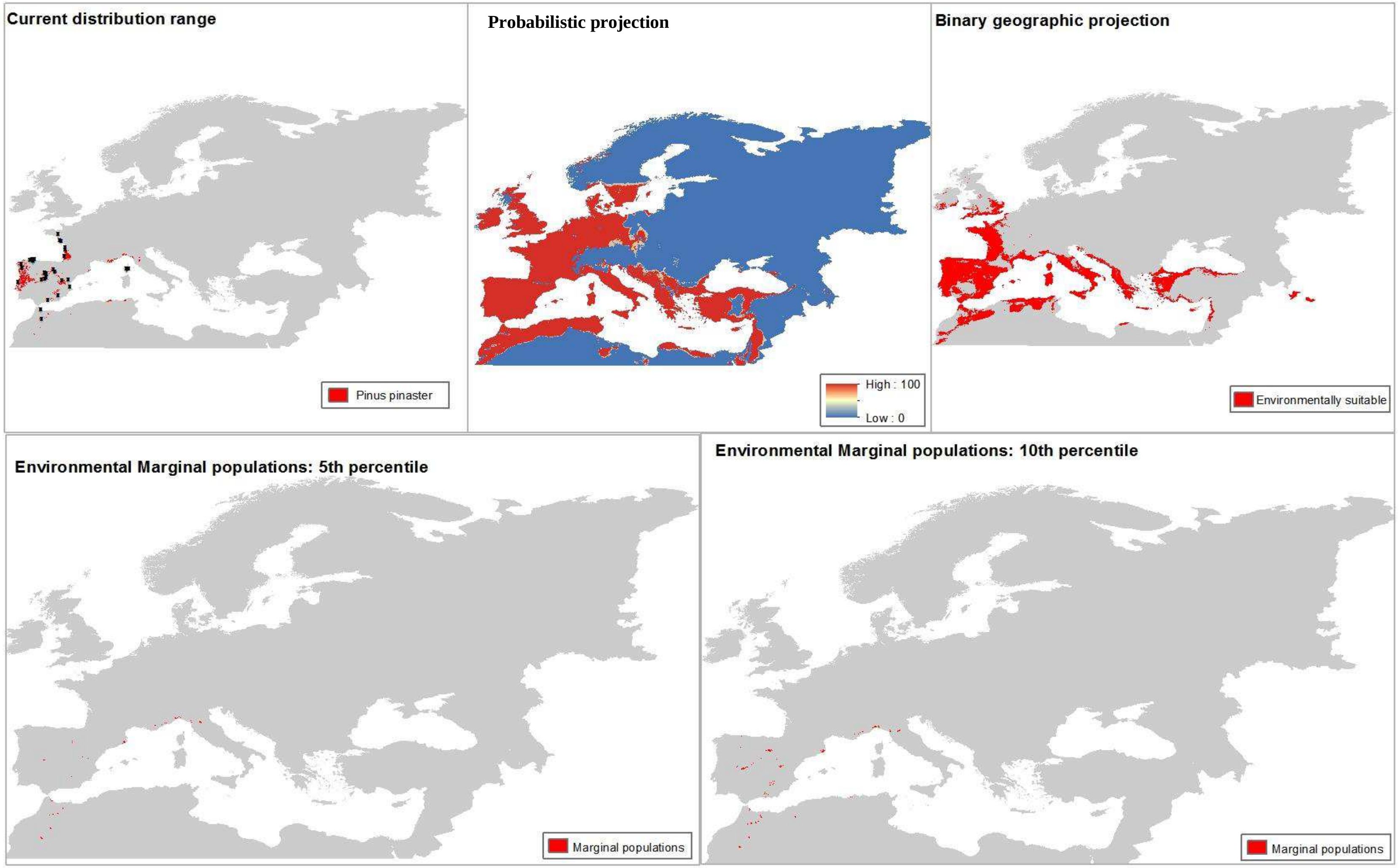
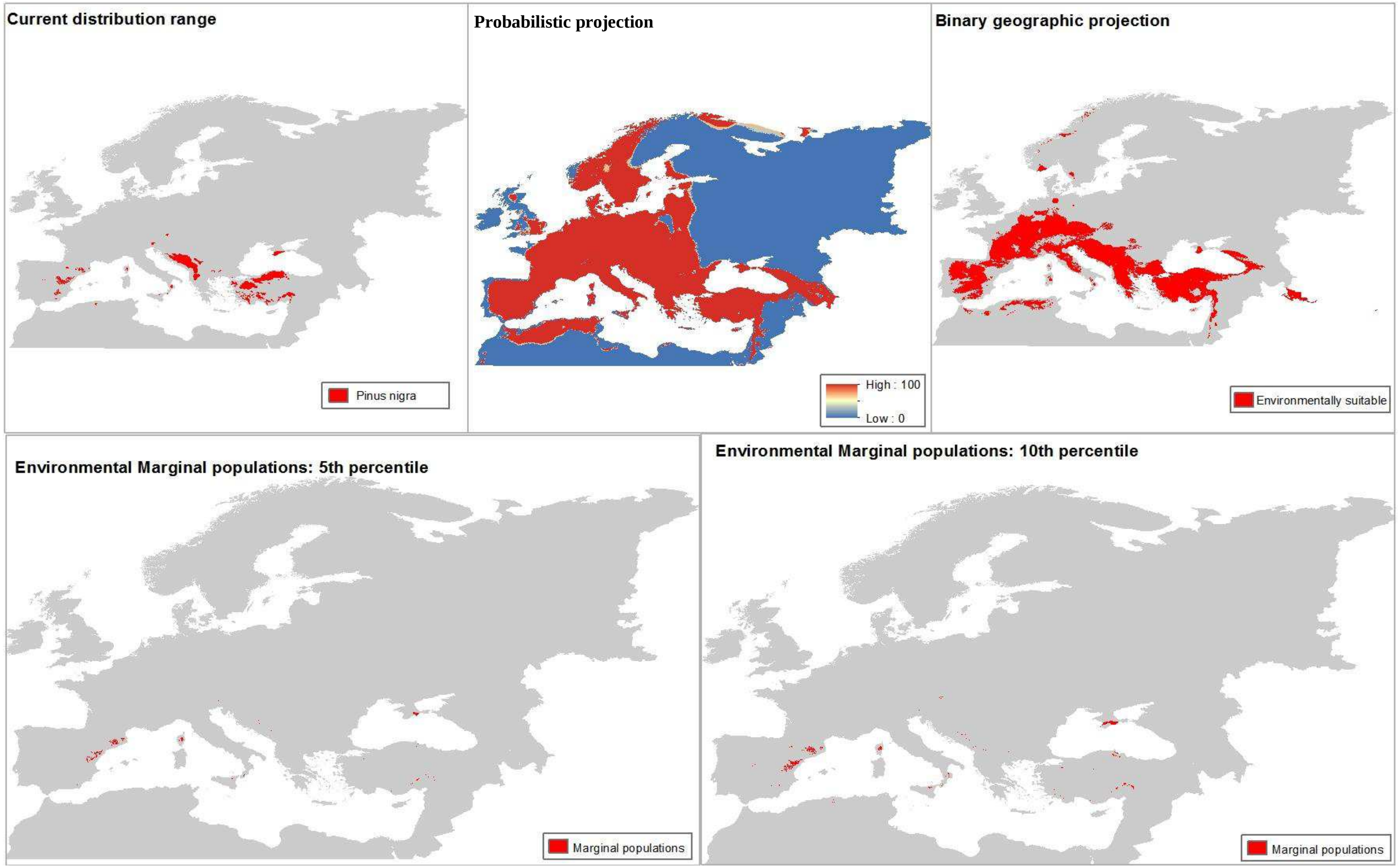


Fig S2: Current distribution, probabilistic and binary geographic projections from Species Distribution Models, and environmental marginal populations (two different thresholds: 5th and 10th percentiles) of *Pinus pinaster* Ait. Black dots correspond to genetically characterized populations.


Current distribution range

Probabilistic projection

Binary geographic projection

Environmental Marginal populations: 5th percentile

Environmental Marginal populations: 10th percentile

Fig S3: Current distribution, probabilistic and binary geographic projections from Species Distribution Models, and environmental marginal populations (two different thresholds: 5th and 10th percentiles) of *Pinus nigra* Arnold. Black dots correspond to genetically characterized populations.

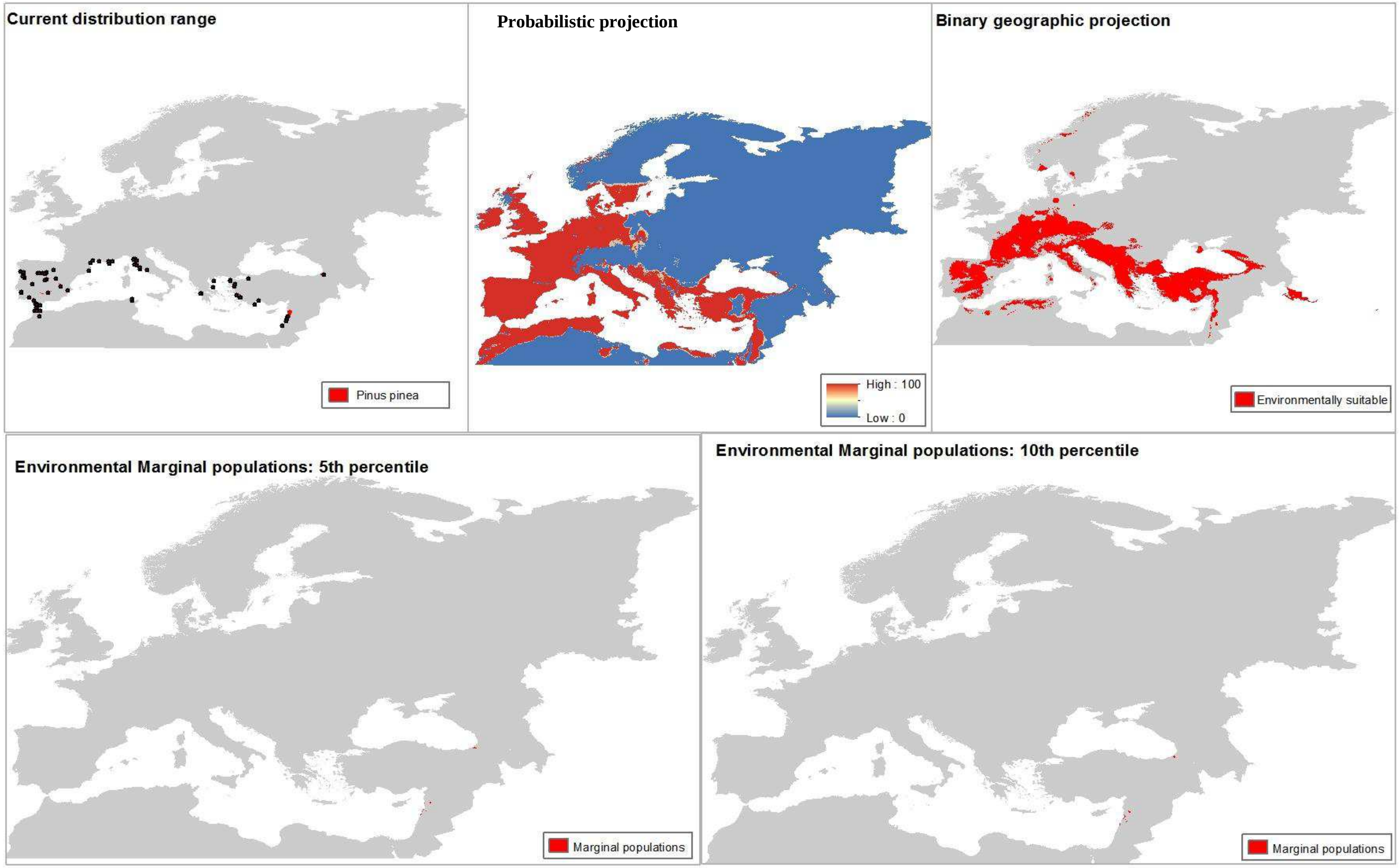


Fig S4: Current distribution, probabilistic and binary geographic projections from Species Distribution Models, and environmental marginal populations (two different thresholds: 5th and 10th percentiles) of *Pinus pinea* L. Black dots correspond to genetically characterized populations.

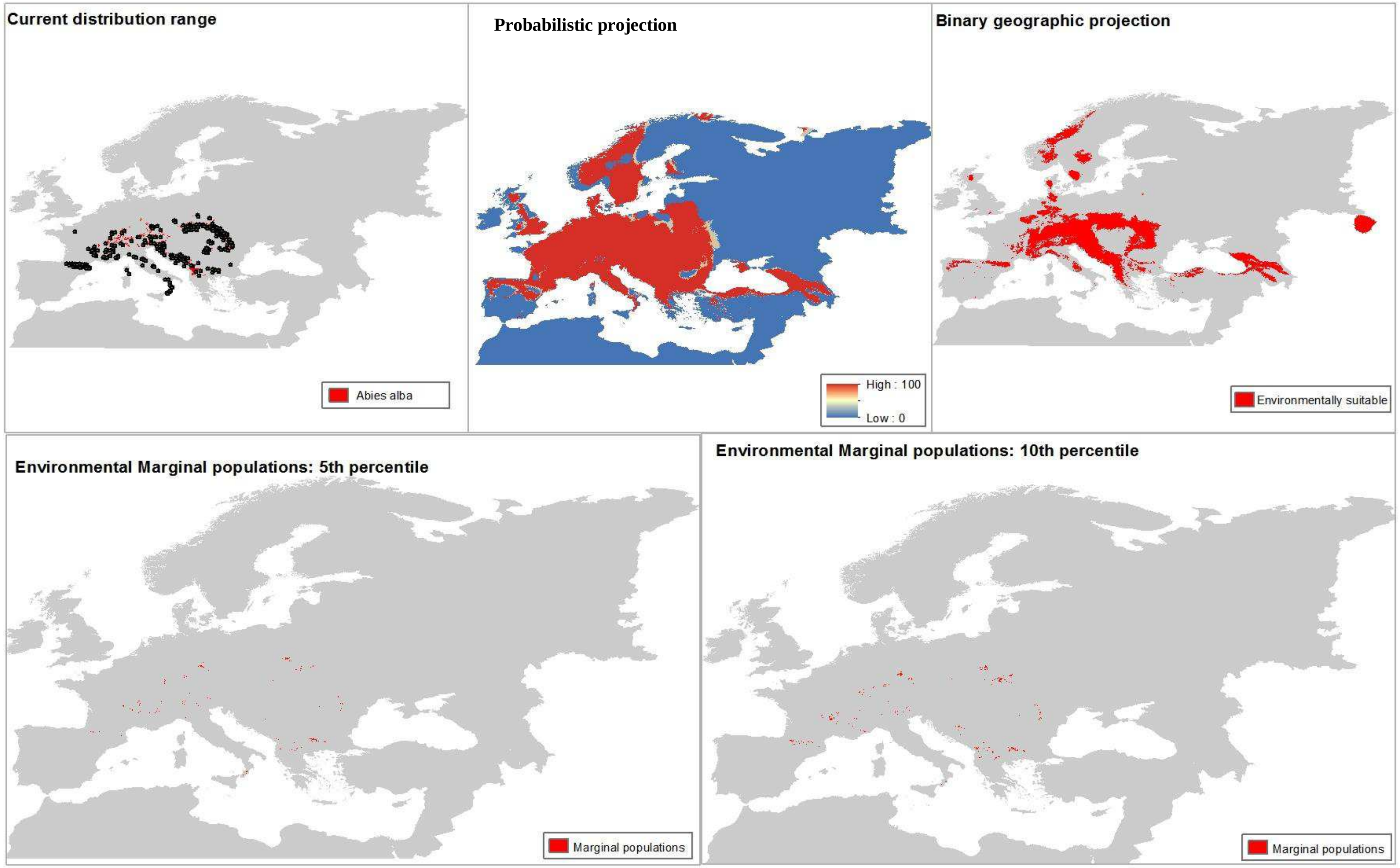


Fig S5: Current distribution, probabilistic and binary geographic projections from Species Distribution Models, and environmental marginal populations (two different thresholds: 5th and 10th percentiles) of *Abies alba* Mill. Black dots correspond to genetically characterized populations.

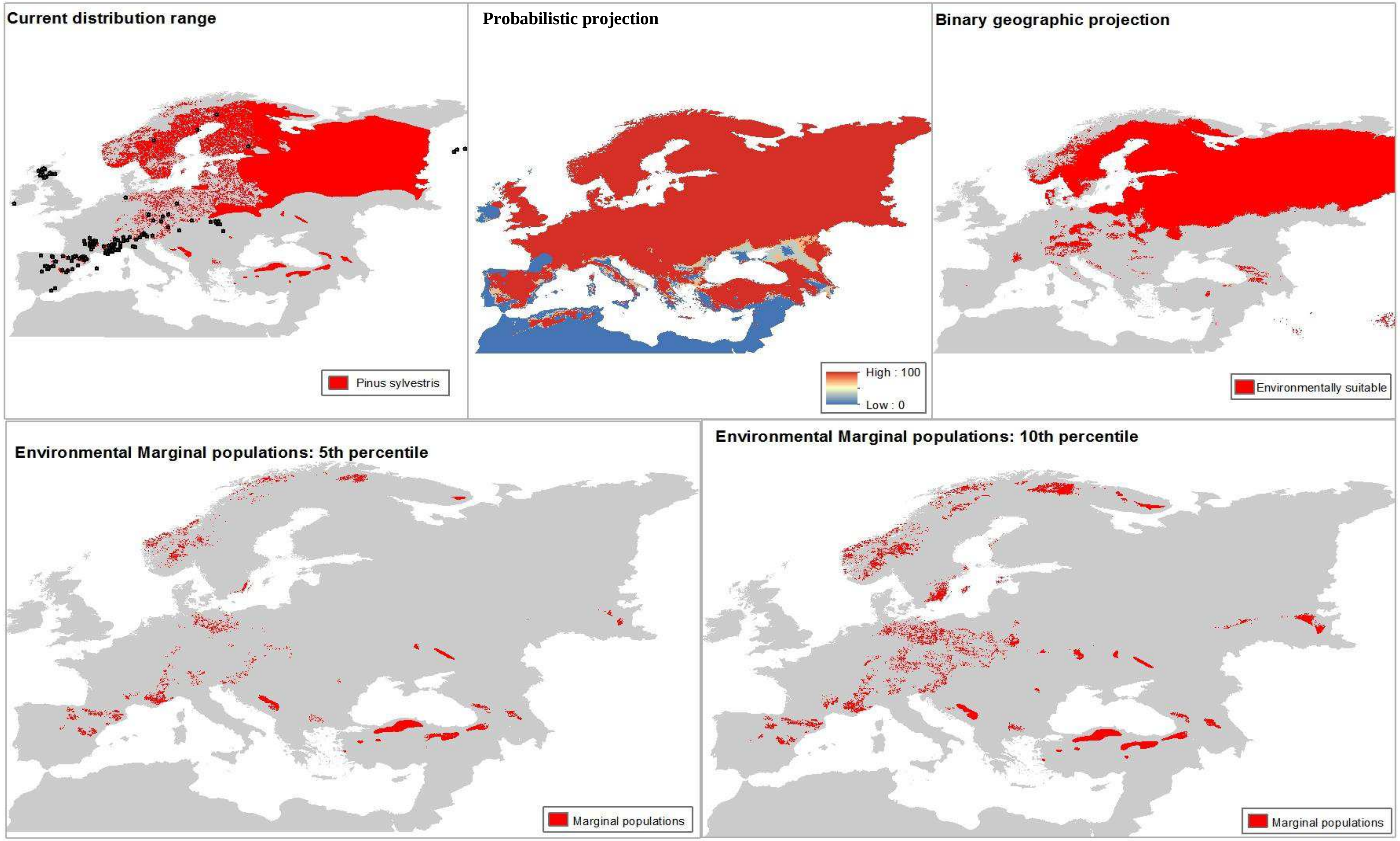


Fig S6: Current distribution, probabilistic and binary geographic projections from Species Distribution Models, and environmental marginal populations (two different thresholds: 5th and 10th percentiles) of *Pinus sylvestris* L. Black dots correspond to genetically characterized populations.

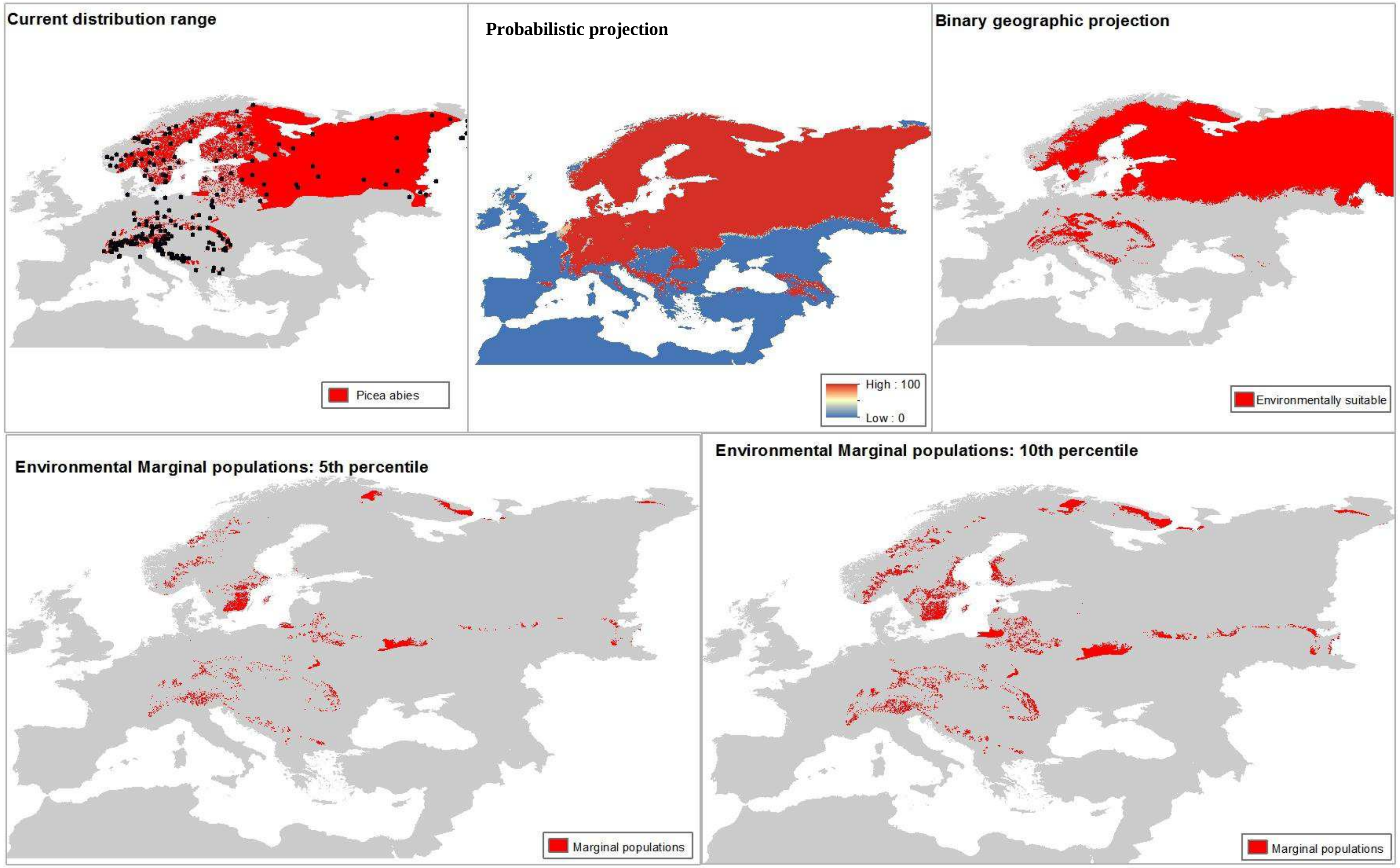


Fig S7: Current distribution, probabilistic and binary geographic projections from Species Distribution Models, and environmental marginal populations (two different thresholds: 5th and 10th percentiles) of *Picea abies* L. Black dots correspond to genetically characterized populations.

APPENDIX S2

Geographic and environmental indices for genetically characterized populations

Table S2: Geographical – distance to the border (D_{border}), size of the core patch ($Size_{core}$), distance to the nearest big (>100 ha) core path ($D_{largepatch}$) - and environmental indexes (Env. Index; corresponding percentile in brackets) for genetically characterized populations of *Pinus halepensis* Mill.

Population	Lat	Long	D_{border} (km)	$Size_{Core}$ (ha)	$D_{largepatch}$ (Km)	Env. Index
Amfilohia	21.28	38.88	-79.54	163025.08	80.72	72.5(7.2)
North Eubea	23.18	38.58	-5.37	21441.77	53.43	94.3(60.4)
Elea	21.53	37.77	-15.53	163025.08	81.50	90.2(32.1)
Zaouia Ifrane	-5.14	33.57	-40.36	131965.81	158.41	87.8(24.6)
Kassandra	23.88	40.09	-139.63	32500.26	140.73	95.7(74)
Imperia	8.05	43.9	-140.07	407490.13	141.03	90.0(31.3)
Shaharia	34.83	31.6	-1179.42	522406.96	1180.24	1.5(0.1)
Litorale Tarantino	17.12	40.62	-473.20	163025.08	474.34	96.5(82.1)
Thala	8.65	35.57	6.27	778425.03	69.52	96.2(79.7)
Aures Beni Melloul	6.83	35.17	-7.30	540120.92	110.35	96.6(83.1)
Quercianella	10.34	43.49	-317.08	407490.13	317.78	87.7(24.4)
Tabarka	9.08	36.51	2.13	242335.82	186.06	93.4(52)
Alcotx	4.17	39.97	2.68	26692.73	67.77	94.6(63)
Carlo Forte	8.18	39.08	-278.66	242335.82	279.81	-
Santanyi	3.05	39.28	-5.18	993.11	52.38	96.1(78.8)
Otricoli	12.38	42.24	-504.89	407490.13	505.76	85.0(18)
Gargano Marzini	15.94	41.9	-638.47	163025.08	639.64	93.5(52.8)
Garzano Monte Pucci	15.86	41.55	-619.59	163025.08	620.75	96.1(78.8)
Nat	35.03	32.72	-1149.90	5946.95	1150.96	8.3(0.2)
Palma de Mallor	2.94	39.15	-20.21	993.11	55.06	96.6(83.1)
Alcudia	3.17	39.87	-8.37	58913.98	51.29	92.7(45.9)
Zuera	-0.92	41.92	1.67	30444.19	53.29	92.0(41.1)
Atalix	4.05	39.92	-6.06	26692.73	56.24	94.3(60.4)
Cabanellas	2.78	42.25	8.79	110454.99	53.51	84.7(17.4)
Tivissa	0.76	41.06	12.62	1070044.22	75.43	93.7(54.6)
Serra d'Irta	0.32	40.35	-9.24	2336787.15	98.16	94.5(62.1)
Benicassim	0.03	40.08	-5.20	2336787.15	51.46	95.6(72.7)
Tuéjar	-1.16	39.82	14.44	2336787.15	50.01	95.3(69.2)
Alhama de Murcia	-1.53	37.86	-2.13	2701.27	51.34	98.3(97.9)
Benamaurel	-2.74	37.7	0.51	1828.14	54.61	95.7(74)

Frigiliana	-3.92	36.82	1.56	162332.66	58.14	95.6(72.7)
Carratraca	-4.83	36.84	3.98	26627.30	76.67	92.7(45.9)
Tibi	-0.65	38.52	4.55	2336787.15	51.81	92.6(45.1)
Serra d'Irta	0.32	40.35	-9.24	2336787.15	98.16	94.5(62.1)
Bicorp	-0.86	39.1	20.40	2336787.15	50.34	98.0(96.3)
Serra Calderona	-0.48	39.74	6.42	2336787.15	66.29	95.9(76.7)
Alcantud	-2.31	40.56	-28.53	195.74	57.41	79.0(10.8)
Colmenar de Oreja	-3.33	40.09	-50.58	781.16	51.74	96.8(85.4)
Santiago de la Espada	-2.47	38.23	12.33	2336787.15	59.29	96.5(82.1)
Serra Calderona	-0.48	39.74	6.42	2336787.15	66.29	95.9(76.7)
Villajoyosa	-0.3	38.5	-11.28	2336787.15	59.99	97.4(92.2)
Monovar	-0.96	38.39	2.00	2336787.15	55.37	98.3(97.9)
Montan	-0.59	40.05	-0.51	2336787.15	65.02	94.0(57.4)
Cabanes	0.04	40.1	7.27	2336787.15	77.92	95.5(71.5)
Sinarcas	-1.2	39.8	13.19	2336787.15	52.01	91.5(38.3)
Titaguas	-1.3	39.89	-0.45	2336787.15	64.27	85.9(19.9)
Cabanes	0.04	40.1	7.27	2336787.15	77.92	95.5(71.5)
Alzira	-0.39	39.12	2.20	49494.60	69.75	96.7(84.2)
Eslida	-0.29	39.87	7.34	2336787.15	78.04	96.1(78.8)

Table S3: Geographical – distance to the border (D_{border}), size of the core patch ($Size_{core}$), distance to the nearest big (>100 ha) core path ($D_{largepatch}$) and cost-distances to the centroid of the distribution ($D_{centroid}$) - and environmental indexes (Env. Index; corresponding percentile in brackets) for genetically characterized populations of *Pinus pinaster* Ait.

Population	Lat	Long	D_{border} (km)	$Size_{Core}$ (ha)	$D_{bigpatch}$ (Km)	$D_{centroid}$	Env. Index
Pleucadec	-2.34	47.78	-233.26	417.63	233.88	13593.57	83.0(15.4)
Hourtin	-1.15	45.18	1.70	1292246.35	54.30	7518.83	92.6(37.9)
Mimizan	-1.30	44.13	-0.38	1292246.35	69.42	6237.69	89.5(26.8)
Olonne sur Mer	-1.83	46.57	-95.21	417.63	95.65	11147.12	92.3(36.6)
Petrocq	-1.30	44.06	1.71	1292246.35	63.22	6145.10	91.5(33.2)
StJean des Monts	-2.03	46.76	-121.93	417.63	122.36	11704.16	92.6(37.9)
Leverdon	-1.09	45.55	-4.31	119.83	52.54	8028.02	94.3(46.2)
Alto de la Llama	-6.49	43.28	-17.31	28337.31	50.62	5429.84	95.2(51.9)
Armayan	-6.46	43.30	-13.65	28337.31	55.31	5412.26	94.4(46.8)
Cadavedo	-6.42	43.54	-11.21	28337.31	94.51	5689.60	87.9(23)
Castropol	-6.98	43.50	-50.99	5543287.95	51.93	6417.39	95.1(51.2)
Lamu±o	-6.22	43.56	-0.44	28337.31	99.71	5477.12	81.4(13.5)
Leiria	-8.96	39.78	3.46	5543287.95	60.72	5603.88	96.8(63.4)
Puerto de Vega	-6.63	43.55	-28.46	28337.31	54.88	6003.76	90.2(28.7)
Rodoiros	-6.54	43.43	-19.38	28337.31	54.52	5699.99	92.3(36.6)
San Cipriano de Ribaterme	-8.36	42.12	28.29	5543287.95	52.35	5701.62	99.1(87.6)
Segurde	-8.45	42.82	26.99	5543287.95	51.21	6274.23	97.1(66.2)
Sierra de Barcia	-6.49	43.53	-17.30	28337.31	88.46	5817.20	89.3(26.3)
Tamrabta	-5.02	33.60	-23.82	21447.27	61.03	16972.04	94.2(45.7)
Boniches	-1.66	39.99	19.85	566441.97	56.03	4471.16	93.0(39.8)
Olba	-0.62	40.17	-41.82	566441.97	65.63	5502.86	95.8(56.4)
San Leonardo	-3.06	41.83	6.31	186452.57	57.56	1328.94	79.8(11.7)
Valquemada	-4.31	40.52	18.16	378431.57	50.62	978.80	94.3(46.2)
Pineta	9.04	41.97	2.06	110436.87	217.86	32795.50	96.6(61.9)
Madisouka	-5.23	35.18	1.26	120547.09	116.52	13540.41	98.2(76.7)
Bayubas de Abajo	-2.88	41.52	1.17	186452.57	50.02	1575.99	92.6(37.9)
Carbonero el Mayor	-4.28	41.17	-1.55	219842.08	52.83	123.84	94.1(45.2)
Coca	-4.50	41.25	4.14	219842.08	60.24	274.61	95.1(51.2)
Arenas de San Pedro	-5.12	40.19	-3.06	378431.57	64.23	1583.05	92(35.3)
Cuellar	-4.48	41.37	12.10	219842.08	51.67	324.89	95.2(51.9)
Oria	-2.35	37.53	-44.93	3986.62	63.52	7886.41	90(28.1)
Cenicientos	-4.49	40.28	8.25	378431.57	60.65	1285.61	90.9(31)
Pinia	9.46	42.02	-8.61	110436.87	188.05	33048.86	79.1(11)
Competa	-3.95	36.83	-37.84	128915.67	153.03	9036.95	80.4(12.4)
Quatretonda	-0.36	38.97	-79.37	566441.97	80.54	6814.63	78.3(10.4)

Table S4: Environmental indexes (corresponding percentile in brackets) for genetically characterized populations of *Pinus pinea* L.

Population	Lat	Long	Env. index
Ponte de Lima	-8.60	41.77	98.9(99.8)
Vieira do Minho	-8.10	41.68	99.8(100.0)
Amarante	-8.10	41.30	99.4(100.0)
Serra do Minho	-8.15	41.00	99.5(100.0)
Viseu Figueira do Campo	-7.90	40.66	99.4(100.0)
Vila Cha de Sa (Viseu)	-7.90	40.67	99.4(100.0)
Alcacer do Sal	-8.52	38.38	83.5(65.9)
Alcácer Sal Herdade Ervideira	-8.50	38.25	79.1(59.6)
Vega Sicilia	-4.30	41.62	91.6(81.0)
Valorio	-5.77	41.52	95.3(95.6)
Toro	-5.45	41.52	95.6(97.8)
Tordesillas	-4.95	41.50	95.4(96.3)
Portillo	-4.52	41.50	92.9(85.0)
Montemayor de Pililla	-4.42	41.50	93.1(86.0)
Iscar	-4.52	41.35	95.3(95.6)
Cogeces de Iscar	-4.52	41.41	94.8(92.8)
Budia	-2.75	40.67	87.8(74.1)
Sistema Central	-4.33	40.50	94.1(90.0)
Cadalso	-4.52	40.28	94.1(90.0)
Garrovillas	-6.58	39.68	70.7(51.1)
Tarazona de la Mancha	-1.92	39.28	94.4(91.1)
Biar	-0.75	38.63	89.5(77.0)
Sierra Morena	-4.00	38.17	11.3(12.0)
Cartaya	-7.18	37.37	55.8(39.7)
Doñana	-6.42	36.92	46.2(33.2)
Conil de la Frontera	-6.08	36.33	86.9(72.3)
Conil de la Frontera	-6.03	36.33	86.8(72.1)
Las Lomas	-5.87	36.30	84.5(67.9)
Mizzine	-5.35	35.10	90.3(78.3)
Cap Spartel	-5.92	35.78	78.8(59.3)
Cap Spartel	-5.89	35.79	84.3(67.5)
Koudia Hamra	-6.17	35.18	63.1(44.4)
Ain Grana	-5.33	35.17	84.6(68.1)
Forêt d'Izarène	-5.47	34.82	73.9(54.0)
Cataluna Litoral 1	2.83	41.83	38.9(26.9)
Parafrugell	3.10	41.95	81.5(62.7)
Vinassan (Aude)	3.08	43.20	78.0(58.4)
Villeneuvette (Hérault)	3.40	43.62	87.0(72.5)
Saintes-Maries (B. du Rh.)	4.42	43.47	82.5(64.3)
Le Val / Brignoles (Var)	6.08	43.43	93.0(85.5)
Hyères (Var)	6.15	43.08	91.7(81.2)
St Aygulf (Var)	6.68	43.43	96.9(99.2)

St Raphaël (Var)	6.77	43.43	97.7(99.5)
S. Rossore Via dei Pini-Via dei Bicchi (Pi)	10.28	43.75	93.3(86.9)
Migliarino	10.28	43.78	93.8(88.8)
Cecina	10.30	43.75	93.4(87.3)
Tromboli di Cecina	10.47	43.32	74.5(54.7)
Tomboli di Cecina Rosignano Bibbona (Pi)	10.53	43.25	75.2(55.4)
Follonica Tombolo settentrionale (Gr)	10.77	42.92	72.5(52.7)
Follonica Bandita Scarlino Gavorrano (Gr)	10.83	42.92	73.2(53.4)
Feniglia duna Orbetello (Gr)	11.28	42.42	67.8(48.3)
Dar Chichou (Cap Bon)	10.98	36.95	22.2(16.2)
Bizerte	9.93	37.25	59.3(42.0)
Metochi Patras	21.50	38.10	89.7(77.3)
Kunupeli	21.37	38.12	85.0(68.8)
Strophillia	21.38	38.13	84.8(68.5)
Mandraki	23.41	39.17	41.8(29.5)
Polygyros (Chalkidiki)	23.68	40.23	15.3(13.6)
Kumluca	30.33	36.30	19.6(15.2)
Serik	31.02	36.87	25.9(18.0)
Mugla / Katranci / Turtuglu	28.05	37.37	63.1(44.4)
Yatagan-Katranci	27.92	37.45	75.5(55.7)
Aydin Karine	27.38	37.77	81.6(62.9)
Izmir / Bergama / Kartai (n° 3227)	26.95	39.20	97.0(99.3)
Balikesir / Edremit / Atkayasi 3226	27.10	39.67	94.2(90.3)
Canakkale / Eceabat / Milli Parki (n°3243)	26.27	40.18	96.1(98.6)
Çanakkale	26.28	40.33	93.7(88.5)
Bursa / Yalova / Dumanlidag (3225)	29.37	40.53	97.7(99.5)
Yalova	29.38	40.54	98.1(99.6)
Artvin 3229	41.85	41.18	88.7(75.6)
Bkassine	35.57	33.55	6.7(10.0)
Kornael	35.72	33.85	87.8(74.1)
Qsaibe	35.68	33.88	63.5(44.7)
Beit Mounzer	35.91	34.29	80.9(61.8)
Monte Carmelo	35.00	32.75	0.3(3.4)

Table S5: Environmental indexes (corresponding percentile in brackets) for genetically characterized populations of *Abies alba* Mill.

Population	Lat	Long	Env. index
FP d'Amélie les Bains	2.72	42.42	65.1(6.2)
FC camparan	0.36	42.83	87.1(16.4)
FC de Sainte Engrâce (bois d'Arbouty)	-0.81	43.01	72.4(8.3)
FD des Corbieres occidentales et FC d'Arques	2.40	42.93	23.5(0.7)
FC d'Aspet	0.48	43.00	83.0(13.4)
FD des 3 Asses	6.23	44.40	84.8(14.6)
FC d'Aulus les Bains	1.39	42.77	93.2(27.2)
FC d'Aragnouet	0.12	42.47	15.6(0.4)
FC Bagnères de Bigorre	0.10	43.00	81.4(12.3)
FD Barousse	0.47	42.93	86.2(15.7)
FP : GF de Cabrefol	1.96	42.87	77.1(10.2)
FD de Bethmale	1.96	42.87	77.1(10.2)
FC de Bielle et Bilhères	-0.50	43.06	66.2(6.5)
FC de Bolquère	2.06	42.53	87.4(16.7)
FC de Borce	0.34	42.54	58.3(4.9)
FD de Boscodon	6.28	44.32	95.8(40.6)
FC de Boutx	0.76	42.91	90.9(21.1)
FC de la Brigue	7.37	44.03	73(8.5)
FD de Cagne	9.44	41.35	-
FD Callong-Mirailles	2.10	42.87	72.8(8.4)
FD de l'Eau salée (Camps sur Agly)	2.42	42.89	41.5(2.1)
FD du Canigou	2.41	42.50	84.0(14.0)
Forêt de la Commission syndicale de la Vallée de Saint-Savin	0.64	42.53	90.1(19.8)
FD de Celles	6.57	48.27	58.1(4.8)
FD de la Grande Chartreuse	5.48	45.20	25.8(0.8)
FD du Canigou	2.25	42.34	85.4(15.0)
FP de Contrazy	1.22	43.07	7.4(0.2)
FD du Mont Lozère	3.60	44.47	49.8(3.2)
FC de Cruis	5.50	44.03	86.2(15.7)
Forêt du Pré de la Dame	3.90	44.39	69.9(7.5)
FD du Donon	7.09	48.29	66.2(6.5)
FD de Belissens	1.35	42.95	81.8(12.6)
FS La Fage et le Réal	3.16	45.04	16.9(0.4)
FD Callong-Mirailles	1.98	42.75	89.9(19.5)
FD des Fanges	1.97	42.77	85.8(15.3)
FC de Gèdre	0.11	42.47	14.4(0.4)
FD de Gerardmer	6.52	48.04	96.6(48.6)
FD de Montnaie-Gravas	2.93	42.42	1.4(0.0)

FD des Hares	2.28	42.83	57.9(4.8)
F syndicale	2.02	42.70	89.1(18.5)
FP : GF d'Hèches	-0.03	42.99	73.6(8.7)
FC d'Iraty	-1.07	43.00	78.2(10.7)
FS d'Issaux	0.50	42.85	87.9(17.2)
FD de la Joux	5.57	46.51	81.1(12.2)
FD de la Cigalère	0.65	42.81	86.1(15.6)
Bois des Laitiers	0.37	48.45	4.4(0.1)
FP : Syndicale d'Assouste	-0.41	42.96	69.3(7.3)
FC de la Léchère Doucy	6.28	45.31	97.5(61.8)
FP : GF de Vilhac	1.54	42.56	92.3(24.2)
FD du Livradois	3.31	45.32	2.8(0.1)
FC de Bagnères de Luchon	0.65	42.72	92.8(25.8)
FD de Lure	6.29	47.41	82.0(12.7)
FD de la Haute-Ariège	1.86	42.67	91.1(21.4)
FP de Montbrun Bocage	1.22	43.11	9.1(0.2)
Forêt du Bois Noir	3.84	44.44	66.1(6.4)
FC de Montferrier	1.78	42.86	92.3(24.2)
FS d'Issaux	0.50	42.95	82.3(12.9)
FC de Planès	2.82	42.29	1.3(0.0)
FD de Prémol	5.48	45.07	97.3(58.2)
FD Punteniellu	9.11	41.99	0.0(0.0)
FD Punteniellu	9.11	41.99	0.0(0.0)
Forêt du Sapet	3.60	44.47	49.8(3.2)
FD du Val de Siguer	1.33	42.46	93.5(28.4)
FC de St Etienne	4.29	45.22	65.8(6.4)
FD St Lary	0.87	42.90	94.8(34.0)
Haut Vallespir	2.34	42.27	67.5(6.8)
FC de Beaumont de ventoux	5.09	44.11	10.1(0.2)
FC de Beaumont de ventoux	5.09	44.11	10.1(0.2)
FD Ste Croix Volvestre (foret ancienne)	1.17	43.11	5.5(0.1)
FD Ste Croix Volvestre (foret récente)	1.16	43.11	3.8(0.1)
Jazero	20.73	48.73	95.6(38.9)
Zwierzyniec	22.97	50.62	34.3(1.4)
Cisna	22.30	49.18	93.7(29.1)
Borynja	22.97	49.08	93.3(27.6)
Rovte-Lavrovec	15.17	45.98	88.8(18.1)
Nyrsko	13.10	49.27	94.3(31.6)
Bily Potok	13.57	49.15	97.4(59.9)
Zofin	14.68	48.65	97.7(66.0)
Pisek	14.15	49.30	84.7(14.5)
Mala Morava	17.02	50.23	91.6(22.4)
Stribrnice	16.85	50.18	89.8(19.3)
Borsucie	19.50	49.57	98.9(92.6)
Hnilcik	20.58	48.85	98.6(85.5)

Polana	19.50	48.65	96.2(44.4)
Dobroc	19.73	48.72	98.8(90.7)
Sitno	18.85	48.37	91.0(21.3)
Boryslav	23.40	49.28	89.6(19.1)
Vorotec	26.07	46.85	98.5(83.2)
Baiut	24.00	47.58	97.2(56.6)
Valea Ierii	23.35	46.65	96.6(48.6)
La Tour d'Auvergne	2.68	45.53	45.7(2.6)
St Haon	3.77	44.85	51.6(3.6)
Mieussy	6.53	46.13	72.9(8.5)
Grande Chartreuse	5.72	45.37	96.5(47.5)
F2	6.66	45.04	97.3(58.2)
F3	6.81	45.61	88.8(18.1)
A3	14.68	47.54	99.6(97.7)
Altensteig	8.62	48.58	93.0(26.4)
Hornberg	8.23	48.22	93.4(28)
Nagold Hornberg-	8.73	48.55	88.1(17.4)
Nazarje	14.52	46.27	98.8(90.7)
Postojna	14.22	45.78	92.6(25.1)
Idrija	14.03	46.00	78.6(10.9)
Pohorje-Hudi Kot	15.27	46.52	98.9(92.6)
Bohor	15.28	46.07	96.6(48.6)
Hrusica	14.13	45.55	76.0(9.7)
Zelezna Ruda	13.23	49.13	97.3(58.2)
Kubova Hut	13.42	49.00	96.1(43.4)
Jirikovo udoli	14.67	48.67	97.7(66.0)
Javorina	20.15	49.27	99.0(94.2)
Stara Voda	20.68	48.80	95.7(39.7)
Vtacnik	18.62	48.62	96.1(43.4)
Hedwizyn	22.82	50.58	40.4(2.0)
Toporcza	22.97	50.65	41.6(2.1)
Dukla	21.68	49.55	88.1(17.4)
Losie	21.07	49.58	93.4(28.0)
Tarnawa Nizna	22.73	49.07	85.9(15.4)
Puscza Biesczads	22.72	49.10	91.2(21.6)
Rozluch	23.33	49.67	69.1(7.3)
Zhdinievo	23.77	48.53	90.7(20.7)
Menczul up	24.37	48.08	97.7(66.0)
Menczul down	24.35	48.07	97.9(70.7)
Jasinja	24.37	48.27	97.6(63.7)
Rarau	25.55	47.47	98.5(83.2)
Valiug-2	22.12	45.13	96.3(45.4)
Retezat	22.75	45.37	97.3(58.2)
Auzelles	3.53	45.62	49.4(3.2)
Forez	3.70	45.93	92.9(26.1)
S1	7.10	46.03	95.1(35.6)
S3	9.87	46.78	79.1(11.2)

D1	10.89	47.54	98.8(90.7)
D2	12.94	47.49	99.3(96.5)
I1	12.13	46.63	51.7(3.6)
Hotedrsica	14.13	45.93	87.4(16.7)
Novo Mesto	15.37	45.80	97.0(53.6)
Kubinska Hola	19.25	49.27	98.7(88.4)
Bumbalka-Salajka	18.42	49.40	97.3(58.2)
Sramkova	18.97	49.17	97.7(66.0)
Brzeznica	19.97	49.28	99.5(97.5)
Biely Vah	20.07	49.10	98.9(92.6)
Cierny Vah	19.90	49.03	98.9(92.6)
Rudno nad Hronom	18.68	48.42	29.5(1.1)
Cervena Skala	20.15	48.82	98.7(88.4)
Mlacik	19.02	48.68	98.4(81.0)
Jasenie	19.45	48.83	97.7(66.0)
Skarzysko-Kam.	20.85	51.10	52.4(3.7)
Nizny Komarnik	21.70	49.38	90.7(20.7)
Nawojowa	20.75	49.53	96.0(42.5)
Moczarna	22.48	49.10	91.8(22.9)
Poiana Brasov	25.57	45.60	98.5(83.2)
Vladeasa	22.78	46.80	97.3(58.2)
Muntele Mic	22.62	45.45	96.5(47.5)
Senj	15.00	44.83	80.0(11.6)
F1	7.27	44.11	86.9(16.3)
A1	13.18	47.13	98.7(88.4)
A2	13.89	47.52	100.0(99.0)
I2	12.10	47.03	96.4(46.4)
Sneznik	14.35	45.58	96.2(44.4)
Velke Karlovice	17.23	50.08	80.4(11.8)
Radim	17.18	50.00	94.8(34.0)
Valaska Dubova	19.30	49.32	99.2(95.8)
Wielka Wies	20.98	50.87	81.7(12.5)
Riaba Skala	22.42	49.10	91.3(21.8)
Moldovita	25.53	47.68	98.6(85.5)
Tazlau	26.62	46.47	45.0(2.5)
Cristian	25.47	45.60	93.5(28.4)
Valiug-1	22.00	45.23	95.4(37.5)
Borovec-Pamp.	23.45	42.15	35.9(1.5)
Jundoly	23.87	42.07	51.4(3.5)
Devin	24.40	41.73	33.9(1.3)
Mavrovo	20.58	41.72	98.5(83.2)
Skopje	21.33	41.83	87.4(16.7)
Dlhy les	20.27	49.22	97.5(61.8)
Stuzica	22.53	49.08	94.2(31.2)
Busteni	25.53	45.45	98.7(88.4)
Poiana Rusca	22.50	44.95	95.9(41.5)
Gospic	15.37	44.53	95.8(40.6)

Delnice	14.63	45.55	90.8(20.9)
Bitola	21.05	41.07	74.0(8.9)
Kopaonik	20.92	43.25	84.1(14.1)
Aspromonte	15.87	38.20	25.7(0.8)
Serra san Bruno	16.30	38.50	20.0(0.5)
Monte Pecoraro	16.17	38.33	10.8(0.3)
Vallombrosa	11.55	43.82	68.4(7.0)
Sila Grande Mucone	16.22	39.47	0.1(0.0)
Listi alto Aspromonte	15.78	38.17	1.8(0.0)
S. Francesco Pollino	16.20	39.90	0.0(0.0)
Serra S. Bruno Pecoraro	16.33	38.55	20.7(0.6)
Camaldoli	11.82	43.80	2.4(0.0)
Campigna	11.73	43.87	0.0(0.0)
Abeti Soprani	14.28	41.87	32.3(1.2)
Serra S. Bruno Archiforo	16.33	38.55	20.7(0.6)
Serra S. Bruno Santa Maria	16.32	38.57	13.3(0.3)
Gariglione	16.45	39.10	22.5(0.7)
Listi Basso	15.78	38.17	1.8(0.0)
Fossa Nardello Aspromonte	15.78	38.17	1.8(0.0)
La Verna	11.92	43.70	63.5(5.8)
Chiusa Pesio	7.65	44.32	52.8(3.8)
Paularo	13.10	46.52	99.0(94.2)
Vrbovsko	15.05	45.40	87.7(17.0)
Skrad	14.88	45.43	94.3(31.6)
Fužine	14.68	45.30	96.2(44.4)
Gerovo	14.60	45.30	70.0(7.5)
Oštrelj	16.38	44.47	98.1(75.2)
Grme?	16.65	44.55	97.0(53.6)
Bos. Grahovo	16.58	44.18	92.7(25.4)
?abulja	17.58	43.53	89.4(18.8)
Biokovo	17.13	43.13	57.3(4.7)
Troglav	16.55	43.95	91.9(23.1)
Glamo?	16.83	44.02	96.0(42.5)
Bugojno	17.30	44.03	97.4(59.9)
Vranica	17.90	43.93	97.7(66.0)
Igman	18.27	43.75	97.9(70.7)
Jahorina	18.57	43.73	94.1(30.8)
Vlaši?	17.48	44.32	96.8(51.0)
Tešanj	18.00	44.57	77.2(10.2)
Zavidovi?i	18.25	44.30	97.4(59.9)
O?evija	18.45	44.17	97.4(59.9)
Klis	18.68	44.10	96.8(51.0)
Knežina	18.75	43.98	96.7(49.8)
Romanija	18.65	43.90	97.6(63.7)
Kalinovik	18.58	43.48	98.4(81.0)
Orjen	18.55	42.63	98.4(81.0)
Bansko	23.39	41.84	51.9(3.6)

Sandansk	23.38	41.65	51.9(3.6)
Kupena	24.30	42.00	54.3(4.1)
Kipilovo	26.28	42.90	60.6(5.3)
Sosogovo	22.55	42.21	51.0(3.4)
Rila	23.37	42.14	52.1(3.7)
Ossau	-0.46	42.86	70.7(7.7)
Lure	5.82	44.11	88.3(17.6)
Vesubie	7.36	43.97	81.3(12.3)
toc	8.46	46.17	97.4(59.9)
sal	6.89	45.05	96.5(47.5)
pes	7.67	44.21	67.2(6.7)
tar	13.60	46.49	99.4(97.2)
noa	11.85	46.13	98.2(77.3)
ner	9.51	44.56	0.0(0.0)
cer	10.24	44.29	0.0(0.0)
abe	10.67	44.14	0(0.0)
btr	12.22	43.60	24.5(0.8)
pig	11.66	42.81	2.0(0.0)
vdc	13.37	42.71	80.2(11.7)
cep	13.44	42.67	72.5(8.3)
cor	13.49	42.62	83.0(13.4)
tos	13.61	42.53	80.8(12.0)
abs	14.29	41.86	32.9(1.3)
cil	15.45	40.40	9.9(0.2)
lau	15.96	40.41	5.6(0.1)
tdp	16.22	39.96	27.2(0.9)
sil	16.64	39.13	56.3(4.5)
ssb	16.35	38.56	30.2(1.1)
gam	15.85	38.14	36.4(1.6)
MNE_Bj	20.01	42.66	96.6(48.6)
MNE_Po	19.77	42.85	98.0(73.0)
MNE_Vst	19.04	43.13	55.1(4.3)
R11	22.25	44.90	94.7(33.5)
R14	21.93	45.08	94.7(33.5)
R22	21.89	45.07	93.7(29.1)
R23	22.48	45.34	96.9(52.2)
R12	22.53	45.39	97.1(55.0)
R15	22.27	45.68	97.1(55.0)
R16	22.46	45.68	97.2(56.6)
R20	24.50	45.64	98.9(92.6)
R21	24.69	45.44	96.7(49.8)
R4	25.53	45.43	98.9(92.6)
R13	25.88	45.28	98.2(77.3)
R2	26.52	45.84	97.0(53.6)
R3	26.73	46.00	79.8(11.5)
R5	26.41	46.65	98.1(75.2)
R6	25.86	46.15	96.5(47.5)

R7	25.19	47.02	98.8(90.7)
R8	26.25	47.17	92.2(23.9)
R9	24.01	47.64	98.4(81.0)
R10	24.62	47.88	97.6(63.7)
R17	25.70	47.82	97.4(59.9)
R18	22.85	46.70	97.8(68.4)
R19	22.62	46.78	97.5(61.8)
R1	23.03	46.46	97.5(61.8)
SRB_Tara	19.31	43.94	97.7(66.0)
F1	7.27	44.11	86.9(16.3)
F2	6.66	45.04	97.3(58.2)
F3	6.81	45.61	88.8(18.1)
F-ISP	5.28	44.18	37.1(1.6)
S1	7.10	46.03	95.1(35.6)
S2	8.05	46.40	80.9(12.1)
S3	9.87	46.78	79.1(11.2)
D1a	10.89	47.54	98.8(90.7)
D1b	11.15	47.51	96.9(52.2)
D2a	12.94	47.49	99.3(96.5)
D2b	12.85	47.67	99.6(97.7)
D-ISP	12.90	47.57	99.8(98.3)
A1	13.18	47.13	98.7(88.4)
A2	13.89	47.52	100.0(99.0)
A2x	13.83	47.48	100.0(99.0)
A3	14.68	47.54	99.6(97.7)
I1	12.13	46.63	51.7(3.6)
I2	12.10	47.03	96.4(46.4)
I-ISP	12.25	46.57	95(35.1)
Kostryna	22.59	48.93	89.2(18.6)
Veretskij pereval	23.17	48.82	95.8(40.6)
Volovets	23.26	48.74	95.9(41.5)
Sojmy	23.49	48.56	95.8(40.6)
Ust' Chorna	23.94	48.34	97.2(56.6)
Lugi	24.53	48.03	97.9(70.7)
Dilove	24.21	47.93	96.9(52.2)
Yasinya	24.32	48.23	97.9(70.7)
Krivopilskij pereval	24.74	48.20	98.0(73.0)
Mykulychyn	24.61	48.42	97.7(66.0)
Beli Oslavi	24.70	48.49	95.6(38.9)
Knyazhdvir	24.90	48.55	92.0(23.4)
Kobyletska Polyana	24.08	48.12	98.2(77.3)
Rosilna	24.40	48.77	93.1(26.8)
Korchivtsi	25.78	47.96	93.3(27.6)
Migove	25.40	48.16	92.5(24.8)
Sepin - Putila	25.29	48.00	98.1(75.2)
Budynets	25.65	48.08	88.4(17.7)
Bagne	25.23	48.22	95.0(35.1)

Rybne	24.61	48.95	88.9(18.2)
Morzhin	23.89	49.15	86.1(15.6)
Rakhinja	24.04	49.02	93.3(27.6)
Tukhla	23.47	48.90	95.8(40.6)
Truskavec	23.51	49.29	84.1(14.1)
Pidbuzh	23.23	49.33	87.3(16.6)
Rozluch	22.97	49.24	93.2(27.2)
Slatioara	25.67	47.47	98.3(79.1)
Campulung Moldovenesc	25.57	47.55	98.3(79.1)
Palotské jedliny NPR	22.03	49.27	95.1(35.6)
Kamenica	21.00	49.21	98.4(81.0)
Baiut Marmaros	24.01	47.66	97.1(55.0)
Lunca Bradului	25.14	47.01	98.6(85.5)
SANT	24.82	47.47	98.2(77.3)

Table S6: Environmental indexes (corresponding percentile in brackets) for genetically characterized populations of *Pinus sylvestris* L.

Population	Lat	Long	Env. index
Abernethy	-3.61	57.21	17.0(2.7)
Allt Cul	-3.35	57.04	9.0(1.2)
Amat	-4.60	57.87	14.0(2)
Ballochbuie	-3.30	56.99	6.8(1)
Beinn Eighe	-5.35	57.63	4.4(0.7)
Black Wood	-4.32	56.67	25.1(4.4)
Coille Coire Chuilc	-4.71	56.41	26.7(4.7)
Cona Glen	-5.33	56.79	21.1(3.6)
Crannach	-4.68	56.58	29.2(5.2)
Glen Affric	-4.92	57.27	16.5(2.5)
Glen Cannich	-4.96	57.33	19.6(3.3)
Glen Derry	-3.58	57.03	19.0(3.1)
Glen Einig	-4.76	57.95	7.9(1.1)
Glen Loy	-5.13	56.91	20.3(3.4)
Glen Tanar	-2.86	57.05	6.2(0.9)
Loch Clair	-5.36	57.56	5.9(0.8)
Meggernie	-4.35	56.58	22.9(4)
Rhidorroch	-4.98	57.89	10.0(1.4)
Rothiemurcus	-3.77	57.15	14.7(2.2)
Sheildaig	-5.64	57.51	2.3(0.4)
Strath Oykel	-4.61	57.98	7.0(1)
lowe Austria east border of the Alps near Pernitz	16.00	47.91	42.3(7.9)
Trevenque (La Cortijuela; southern Spain)	-3.55	37.10	0.9(0.2)
Valsain	4.04	40.87	-
Punkaharju	29.39	61.76	95.2(71)
Kolari	24.05	67.18	84.2(35.5)
Jarocin	17.48	51.97	32.6(5.8)
Krp. Tjärnbergsheden (8 famil)	20.80	64.62	67.8(19.3)
Väster Mj?ingenn (7 famil)	13.58	62.75	76.0(24.5)
Cella di Palmia	10.17	44.63	0.5(0.1)
Dale Wood	-9.63	51.97	0.0(0.0)
Rockforest	8.98	53.00	34.5(6.2)
Vezzano sul Crostolo	10.52	44.52	3.2(0.5)
Carpe	7.78	43.92	1.1(0.2)
Carnino Briga Alta	7.73	44.13	19.8(3.3)
Bossolasco	8.05	44.53	8.3(1.1)
Vezza d'Alba	7.98	44.75	0.2(0.0)
Fenestrelle	7.05	45.03	27.6(4.9)
Savoulx	6.80	45.05	40.6(7.5)

Passerano Marmorito	8.02	45.03	0.4(0.1)
Casalborgone	7.98	45.13	1.3(0.2)
Sarre	7.27	45.75	18.1(2.9)
Toceno	8.48	46.15	63.5(17)
Ticino	8.53	45.53	3.4(0.5)
Olgelasca	9.18	45.73	2.2(0.4)
Val Masino	9.62	46.15	4.8(0.7)
Valvestino Garda	10.58	45.75	14.5(2.1)
Valda	11.27	46.20	36.9(6.6)
Alpe di Siusi	11.58	46.53	56.0(12.8)
Brunico	11.97	46.80	58.2(14.1)
Cortina	12.17	46.53	8.2(1.1)
Claut	12.50	46.27	42.9(8)
Val Dogna	13.32	46.43	61.3(15.8)
Morgex	7.08	45.83	3.2(0.5)
Challand St. Anselme	7.67	45.67	12.6(1.8)
Vysoké Chvojno	15.98	50.12	43.3(8.1)
Zámecký	14.84	48.99	46.3(8.9)
Hadce u Želivky	15.11	49.69	44.1(8.3)
Pluh?v Bor	12.73	50.03	64.3(17.4)
Vl?ek	12.78	50.06	71.0(21.3)
Šumava Dra?í skály	13.48	49.12	43.7(8.2)
Siberia Shapsha	69.46	61.07	-
Siberia Kondinsky Ozera - Green Moss Forest	63.58	60.86	88.7(45.9)
Siberia Kondinsky Ozera - Lichen Forest	63.54	60.86	88.1(44.1)
Siberia Kondinsky Ozera - Ryam Forest	63.52	60.85	88(43.8)
Siberia Mukhrino	68.70	60.89	-
Siberia Khanty-Mansiysk	69.01	60.98	-!
Siberia roadside 1	65.47	61.39	89.1(47.1)
Siberia roadside 2	64.12	61.21	90.9(52.5)
BAZ	-2.85	37.37	30.2(5.4)
BR	-0.58	42.70	28.7(5.1)
BRA	-3.92	41.03	17.6(2.8)
CAM	-3.20	41.22	21.7(3.7)
CERC	-4.06	40.77	26.3(4.7)
COC	-4.50	41.20	4.7(0.7)
COV	-2.82	41.95	16.1(2.5)
CUE	-4.22	41.27	4.2(0.6)
G	-0.68	40.42	17.6(2.8)
GAL	-3.12	41.25	15.8(2.4)
IRE	-2.74	42.02	24.2(4.2)
JAV	-1.00	40.13	12.5(1.8)
LC	0.05	40.75	0.8(0.2)
LILL	-5.25	43.07	0.0(0.0)

LOZ	-3.82	40.98	24.9(4.4)
LLAV	1.21	42.50	27.1(4.8)
MED	-3.30	42.98	4.8(0.7)
MP	1.01	41.34	1.2(0.2)
MU	-1.63	40.52	27.7(4.9)
NAR	-5.12	40.35	36.7(6.6)
NAV	-3.83	41.00	42.4(7.9)
NDA	-4.02	40.76	25.3(4.5)
PEG	-4.20	40.65	29.7(5.3)
PLIL	1.97	42.23	20.2(3.4)
PM	-0.10	42.20	5.2(0.8)
PN	-1.54	42.89	13.3(1.9)
TRE	-3.47	37.09	34.3(6.2)
VAL	-4.02	40.82	20.6(3.5)
ZAD	-3.18	42.83	6.6(0.9)
Cauterets	0.14	42.85	0.3(0.1)
Saint Lary Soulan	0.16	42.83	3.1(0.5)
Saint Lary Soulan	0.31	42.74	20.2(3.4)
Camparan	0.36	42.84	21.7(3.7)
Camparan	0.36	42.82	17.4(2.7)
Marignac	0.67	42.91	15.0(2.2)
Marignac	0.66	42.88	32.1(5.7)
Laruns	0.43	42.91	14.3(2.1)
Lescun	0.42	42.90	20.3(3.4)
Le Port	1.69	42.80	20.1(3.4)
Joucou	2.09	42.81	15.6(2.4)
Galinagues	2.05	42.81	15.8(2.4)
Prades	1.86	42.81	21.3(3.6)
Aston	1.63	42.80	19.8(3.3)
Mérens-les-Vals	1.84	42.63	17.3(2.7)
Les Angles	2.11	42.60	17.6(2.8)
Les Angles	2.11	42.60	17.6(2.8)
Fontrabiouse	2.09	42.63	16.2(2.5)
Osséja	1.98	42.39	18.4(3)
Prats de Mollo	2.38	42.43	10.3(1.4)
Jujols	2.29	42.59	17.2(2.7)
Serdinya	2.30	42.59	15.4(2.3)
Ayguatébia-Talau	2.21	42.54	17.1(2.7)
Allanches	2.93	45.20	69.6(20.4)
Joursac	3.03	45.18	71.2(21.4)
Lieutadès	2.92	44.82	56.2(12.9)
Chaudes-Aigues	3.00	44.81	62.9(16.7)
Ruynes en Margeride	3.26	45.02	53.6(11.5)
Rouffiac	2.17	45.01	0.0(0.0)
Leyvaux	3.09	45.34	76.1(24.5)
Davignac	2.04	45.51	0.0(0.0)
Retournac	3.98	45.23	61.1(15.7)

Le Vernet	3.65	45.04	56.5(13.1)
Vernassal	3.72	45.16	54.7(12.1)
Charraix et St Julien des Chazes	3.55	45.05	53.1(11.2)
St-Pal de Chalencon	3.95	45.37	63.5(17)
Rimeize	3.30	44.74	57.7(13.8)
Chanac	3.38	44.46	15.1(2.2)
Chanac	3.70	45.40	57.1(13.5)
Arlanc	3.64	45.50	58.1(14.1)
Chambon sur Dolore	3.04	45.40	74.2(23.2)
Mazoires	3.06	45.40	75.3(24.0)
Mazoires	3.65	45.53	52.7(11.0)
Le Monestier	2.90	46.11	33.6(6.0)
Menat	3.93	45.57	45.8(8.8)
Saint Antheme	3.93	45.57	45.8(8.8)
Millau	3.18	44.13	13.8(2.0)
Fontan	7.51	44.02	15.8(2.4)
La Brigue	7.64	44.03	11.2(1.6)
St Martin Vésubie	7.26	44.09	16.6(2.6)
Châteauneuf d'Entraunes	6.84	44.15	16.8(2.6)
Guillaumes	6.88	44.03	13.2(1.9)
Château Ville Vieille	6.77	44.72	40.6(7.5)
Château Ville Vieille	6.78	44.73	28.1(5.0)
Montgenèvre	6.69	44.92	40.3(7.4)
Les Vigneaux	6.55	44.83	41.7(7.7)
Mimet	5.49	43.41	2.1(0.4)
Venelles	5.52	43.59	1.1(0.2)
Montfuron	5.68	43.84	2.9(0.5)
Mazaugues	5.86	43.34	2.9(0.5)
Moissac-Bellevue	6.20	43.67	3.3(0.5)
Estoublon	6.22	43.94	3.1(0.5)
Tourrettes	6.72	43.64	0.6(0.1)
Séranon	6.70	43.76	12.5(1.8)
Espenel	5.25	44.67	19.9(3.3)
Le Fugeret	6.67	44.02	20.0(3.3)
Verdaches	6.34	44.24	32.5(5.8)
Bayons	6.18	44.31	18.3(3.0)
St Eusèbe en Champsaur	6.01	44.73	28.7(5.1)
Bakony mountains Feny?f?i ?sfenyves	17.77	47.35	7.0(1.0)
Tinovul Poiana Stampei	25.12	47.30	24.4(4.3)
Turková	19.91	49.02	75(23.8)
Horhanska	24.25	48.43	56.4(13.0)
Holyatunska	23.00	48.77	44.2(8.4)
Mshanska	23.95	48.65	78.0(26.2)
Jaremchanska	24.33	48.40	64.6(17.6)
Zelenska	24.33	48.37	55.7(12.6)
Tatarovska	24.60	48.37	55.1(12.3)

Sokol	24.20	48.70	61.9(16.2)
Mizunska	23.88	48.93	44.4(8.4)
Wuhodska	23.90	48.93	43.0(8.1)
Turova Datcha	24.42	48.88	34.7(6.2)

Table S7: Environmental indexes (corresponding percentile in brackets) for genetically characterized populations of *Picea abies* L.

Population	Lat	Long	Env. index
BIJ-nSSR_1	18.50	44.09	43.7(4.3)
BUG-nSSR_2	17.69	43.96	28.2(2.0)
BUS-nSSR_3	16.82	44.03	67.3(11.0)
DUB-nSSR_4	18.34	44.27	59.2(8.2)
GRM-nSSR_5	16.62	44.66	35.4(2.9)
IGMA-nSSR_6	18.27	43.74	79.2(18.6)
IGMB-nSSR_7	18.27	43.75	79.5(18.9)
PIJ-nSSR_8	18.89	44.04	63.1(9.4)
PRE-nSSR_9	16.49	44.61	73.2(14.0)
RAS-nSSR_10	17.22	44.06	82.2(21.7)
ROM-nSSR_11	18.66	43.90	78.4(17.8)
VLA-nSSR_12	17.45	44.32	66.8(10.8)
ZEL-nSSR_13	18.53	43.49	58.7(8.0)
DOL-nSSR_14	14.37	46.40	84.1(24.0)
LIP-nSSR_15	13.93	46.37	84.3(24.3)
JER-nSSR_16	13.96	46.36	87.2(28.7)
SIJ-nSSR_17	13.99	46.32	88.0(30.1)
CJ-nSSR_18	15.44	46.44	93.7(43.4)
FIS-nSSR_19	15.42	46.50	94.3(45.0)
KON-nSSR_20	15.29	46.42	91.1(37.4)
PEV-nSSR_21	14.01	45.96	37.6(3.2)
HRU-nSSR_22	14.14	45.86	35.4(2.9)
SD-nSSR_23	13.87	45.98	86.0(26.7)
LG-nSSR_24	14.39	45.62	68.9(11.7)
VP-nSSR_25	14.44	45.55	87.6(29.4)
1-nSSR_26	6.94	48.07	66.6(10.7)
2-nSSR_27	7.61	46.18	62.3(9.2)
3-nSSR_28	8.84	46.98	95.2(47.9)
4-nSSR_29	10.59	47.30	32.6(2.5)
5-nSSR_30	11.70	48.00	68.2(11.4)
6-nSSR_31	12.12	47.22	63.2(9.5)
7-nSSR_32	10.71	46.50	7.6(0.4)
8-nSSR_33	12.50	47.00	58.3(7.9)
9-nSSR_34	13.88	47.17	93.7(43.4)
10-nSSR_35	12.30	49.50	55.0(7.0)
11-nSSR_36	13.40	48.90	80.1(19.5)
12-nSSR_37	14.50	49.30	22.2(1.4)
13-nSSR_38	15.50	48.60	31.3(2.4)
14-nSSR_39	15.00	44.70	77.9(17.4)
15-nSSR_40	19.41	43.92	71.7(13.1)
16-nSSR_41	18.10	49.10	63.3(9.5)
17-nSSR_42	21.20	49.00	8.4(0.5)

18-nSSR_43	24.30	48.30	74.1(14.6)
19-nSSR_44	25.90	46.80	23.7(1.5)
20-nSSR_45	25.50	45.30	49.7(5.7)
21-nSSR_46	20.00	50.80	38.5(3.4)
22-nSSR_47	22.96	50.59	26.9(1.9)
23-nSSR_48	5.76	60.82	9.9(0.5)
24-nSSR_49	6.54	60.60	48.6(5.4)
25-nSSR_50	7.34	61.26	-
26-nSSR_51	8.30	58.78	73.5(14.2)
27-nSSR_52	9.88	61.19	57.2(7.6)
28-nSSR_53	10.52	60.09	87.5(29.2)
29-nSSR_54	15.60	57.76	39.3(3.5)
30-nSSR_55	17.74	59.94	53.9(6.7)
31-nSSR_56	13.20	61.40	87.0(28.3)
32-nSSR_57	12.85	61.55	66.9(10.9)
33-nSSR_58	12.43	63.24	70.7(12.6)
34-nSSR_59	12.38	63.29	77.6(17.1)
35-nSSR_60	11.82	63.37	82.3(21.8)
36-nSSR_61	12.73	63.44	79.7(19.1)
37-nSSR_62	12.49	63.49	75.7(15.7)
38-nSSR_63	12.28	63.57	68.8(11.7)
39-nSSR_64	10.96	63.78	51.3(6.1)
40-nSSR_65	15.47	64.67	62.7(9.3)
41-nSSR_66	16.13	65.56	50.3(5.8)
42-nSSR_67	19.92	66.70	83.7(23.5)
43-nSSR_68	27.42	68.40	46.9(5.0)
44-nSSR_69	27.80	65.89	92.9(41.5)
45-nSSR_70	28.05	64.66	98.1(62.1)
46-nSSR_71	29.81	63.10	98.7(68.0)
47-nSSR_72	24.65	62.00	82.6(22.2)
48-nSSR_73	30.03	60.18	95.0(47.2)
49-nSSR_74	26.51	58.06	94.3(45.0)
50-nSSR_75	30.00	57.83	88.9(32.0)
51-nSSR_76	24.05	54.51	51.8(6.2)
52-nSSR_77	31.25	53.50	49.7(5.7)
53-nSSR_78	32.33	54.50	86.9(28.2)
54-nSSR_79	37.50	55.70	95.6(49.3)
55-nSSR_80	34.42	62.88	96.6(53.2)
56-nSSR_81	36.72	62.25	98.1(62.1)
57-nSSR_82	40.00	64.58	97.7(59.0)
58-nSSR_83	40.00	59.33	98.5(65.9)
59-nSSR_84	41.00	57.58	94.3(45.0)
60-nSSR_85	48.58	57.10	98.9(70.1)
61-nSSR_86	52.17	56.25	98.5(65.9)
62-nSSR_87	56.13	54.23	10.7(0.6)
63-nSSR_88	54.17	58.50	99.8(81.7)
64-nSSR_89	54.07	63.95	99.9(86.5)

I19-nSSR_90	11.56	47.32	41.6(3.9)
Q14-nSSR_91	15.08	47.08	29.8(2.2)
R06-nSSR_92	15.20	47.32	94.8(46.5)
R13-nSSR_93	15.53	47.28	64.4(9.9)
S10-nSSR_94	16.02	48.08	12.4(0.7)
Std-nSSR_95	13.29	48.04	77.8(17.3)
X05-nSSR_96	14.64	48.58	86.7(27.8)
Y18-nSSR_97	16.04	47.28	6.2(0.4)
F07-nSSR_98	14.75	47.07	97(54.9)
Gusswerk-nSSR_99	15.35	47.66	92.4(40.4)
Pa22-nSSR_100	60.60	56.80	96.0(50.7)
Pa46-nSSR_101	48.58	57.10	98.9(70.1)
Pa41-nSSR_102	41.00	57.58	94.3(45.0)
Pa45-nSSR_103	40.00	64.58	97.7(59.0)
Pa77-nSSR_104	36.72	62.25	98.1(62.1)
Pa78-nSSR_105	34.42	62.88	96.6(53.2)
Pa44-nSSR_106	30.03	60.18	95.0(47.2)
Pa37-nSSR_107	30.00	57.83	88.9(32.0)
Pa40-nSSR_108	31.25	53.50	49.7(5.7)
Pa79-nSSR_109	24.05	54.51	51.8(6.2)
Pa81-nSSR_110	26.51	58.06	94.3(45.0)
Pa80-nSSR_111	24.65	62.00	82.6(22.2)
Pa25-nSSR_112	29.81	63.10	99.2(73.0)
Pa30-nSSR_113	28.05	64.66	98.2(62.9)
Pa29-nSSR_114	27.80	65.89	92.9(41.5)
FIN-nSSR_115	27.42	68.40	46.9(5.0)
Pa26-nSSR_116	30.07	69.45	28.2(2.0)
Pa33-nSSR_117	19.92	66.70	83.7(23.5)
Pa38-nSSR_118	17.33	66.00	72.8(13.8)
Pa34-nSSR_119	16.13	65.56	55.5(7.2)
Pa42-nSSR_120	16.31	64.79	68.9(11.7)
Pa43-nSSR_121	15.47	64.67	61.3(8.8)
Pa39-nSSR_122	13.20	61.40	87.0(28.3)
Pa83-nSSR_123	17.10	60.80	47.3(5.1)
Pa28-nSSR_124	15.20	60.33	84.7(24.8)
Pa32-nSSR_125	17.74	59.94	53.9(6.7)
Pa82-nSSR_126	15.60	57.76	39.3(3.5)
Pa16-nSSR_127	10.52	60.09	87.5(29.2)
Pa67-nSSR_128	9.88	61.19	59.4(8.2)
Pa11-nSSR_129	8.91	61.09	52.6(6.4)
Pa02-nSSR_130	8.11	60.36	31.8(2.4)
Pa09-nSSR_131	7.42	59.30	60.5(8.6)
Pa18-nSSR_132	8.30	58.78	73.5(14.2)
Pa04-nSSR_133	5.76	60.82	9.9(0.5)
Pa06-nSSR_134	6.34	59.45	33.9(2.7)
Pa05-nSSR_135	6.54	60.60	48.6(5.4)
Pa12-nSSR_136	7.34	61.26	-

F1-nSSR_137	6.94	48.07	66.6(10.7)
CH17-nSSR_138	7.61	46.18	62.3(9.2)
BOE-nSSR_139	8.84	46.98	95.2(47.9)
A14-nSSR_140	10.59	47.30	32.6(2.5)
IU950-nSSR_141	11.70	48.00	68.2(11.4)
A7-nSSR_142	12.12	47.22	63.2(9.5)
I1-nSSR_143	10.71	46.50	7.6(0.4)
IU11-nSSR_144	12.50	47.00	58.3(7.9)
A1-nSSR_145	13.88	47.17	93.7(43.4)
IU22-nSSR_146	12.30	49.50	55.0(7.0)
657-nSSR_147	13.40	48.90	80.1(19.5)
235-nSSR_148	14.50	49.30	22.2(1.4)
625-nSSR_149	15.50	48.60	31.3(2.4)
863-nSSR_150	15.00	44.70	77.9(17.4)
Jug7-nSSR_151	19.41	43.92	71.7(13.1)
1002-nSSR_152	18.10	49.10	63.3(9.5)
1017-nSSR_153	21.20	49.00	8.4(0.5)
1062-nSSR_154	24.30	48.30	74.1(14.6)
1056-nSSR_155	25.90	46.80	23.7(1.5)
1031-nSSR_156	25.50	45.30	49.7(5.7)
IU1005-nSSR_157	20.00	50.80	38.5(3.4)
PL1-nSSR_158	22.96	50.59	26.9(1.9)
Pa041-nSSR_159	5.76	60.82	9.9(0.5)
Pa05-nSSR_160	6.54	60.60	48.6(5.4)
Pa12_3-nSSR_161	7.34	61.26	-
Pa67-nSSR_162	9.88	61.19	57.2(7.6)
Pa320-nSSR_163	17.74	59.94	53.9(6.7)
F-nSSR_164	12.85	61.55	66.9(10.9)
Ho-nSSR_165	12.43	63.24	70.7(12.6)
AS-nSSR_166	12.38	63.29	77.6(17.1)
RT-nSSR_167	11.82	63.37	82.3(21.8)
TN-nSSR_168	12.73	63.44	79.7(19.1)
KL1-nSSR_169	12.49	63.49	75.7(15.7)
MB-nSSR_170	12.28	63.57	68.8(11.7)
MV-nSSR_171	10.96	63.78	51.3(6.1)
43-nSSR_172	15.47	64.67	62.7(9.3)
Pa34-nSSR_173	16.13	65.56	50.3(5.8)
Pa29_9-nSSR_174	27.80	65.89	92.9(41.5)
Pa300-nSSR_175	28.05	64.66	98.1(62.1)
Pa81-nSSR_176	26.51	58.06	94.3(45.0)
Rus4-nSSR_177	32.33	54.50	86.9(28.2)
RUS50-nSSR_178	37.50	55.70	95.6(49.3)
Pa78-nSSR_179	34.42	62.88	96.6(53.2)
SUA-nSSR_180	40.00	59.33	98.5(65.9)
Rus1-nSSR_181	52.17	56.25	98.5(65.9)
SG14-nSSR_182	56.13	54.23	10.7(0.6)
Rus2-nSSR_183	54.17	58.50	99.8(81.7)

KM1-nSSR_184	54.07	63.95	99.9(86.5)
NM10-nSSR_185	49.88	67.27	77.6(17.1)
72_RUS60-nSSR_186	59.87	67.82	84.9(25.1)
70_RUS4-nSSR_187	59.37	61.82	99.8(81.7)
SL-nSSR_188	60.59	56.84	96.0(50.7)
68_RUS1-nSSR_189	58.00	55.00	86.3(27.2)
69_RUS3-nSSR_190	58.87	54.55	99.0(71.1)
SEIDA-nSSR_191	62.88	67.05	93.8(43.6)
KK10-nSSR_192	65.76	66.91	84.2(24.1)
69_RUS2-nSSR_193	65.75	66.90	95.4(48.6)
71_RUS5-nSSR_194	66.47	66.67	99.7(79.4)
Pit-nSSR_195	65.91	65.91	99.8(81.7)
KAZ-nSSR_196	65.61	64.70	99.9(86.5)
Bereva-nSSR_197	65.03	63.93	100.0(99.2)
BER-nSSR_198	64.79	63.99	100.0(99.2)
Oc-nSSR_199	66.09	62.45	99.9(86.5)
XM-nSSR_200	69.25	61.05	-
TOB-nSSR_201	68.37	58.17	-
P01-cpSSR_01	5.33	45.11	0.0(0.0)
P02-cpSSR_02	6.37	46.02	67.2(11.0)
P07-cpSSR_03	10.30	51.40	23.4(1.5)
P08-cpSSR_04	14.00	53.30	5.9(0.3)
P09-cpSSR_05	14.10	51.54	1.9(0.2)
P10-cpSSR_06	12.30	50.08	42.4(4.1)
P11-cpSSR_07	10.41	50.42	3.6(0.2)
P15-cpSSR_08	8.40	46.66	4.4(0.3)
P19-cpSSR_09	11.20	49.54	9.6(0.5)
P20-cpSSR_10	13.25	48.33	7.6(0.4)
P23-cpSSR_11	10.49	47.50	73.2(14)
P28-cpSSR_12	11.54	47.29	3.4(0.2)
P32-cpSSR_13	15.21	47.10	57.0(7.6)
P37-cpSSR_14	13.20	49.30	73.7(14.3)
P38-cpSSR_15	14.20	50.65	78.5(17.9)
P39-cpSSR_16	15.37	50.87	85.2(25.5)
P40-cpSSR_17	14.05	49.22	34.8(2.8)
P43-cpSSR_18	17.18	49.45	1.9(0.2)
P48-cpSSR_19	20.09	49.13	97.9(60.6)
P54-cpSSR_20	16.12	44.75	79.7(19.1)
P55-cpSSR_21	20.52	42.02	0.0(0.0)
P56-cpSSR_22	25.00	42.13	0.0(0.0)
P57-cpSSR_23	25.17	45.36	11.8(0.7)
P58-cpSSR_24	23.47	46.35	27.0(1.9)
P59-cpSSR_25	25.00	47.20	57.0(7.6)
P60-cpSSR_26	22.50	49.03	65.9(10.5)
P61-cpSSR_27	21.15	51.15	18.4(1.1)
P64-cpSSR_28	16.41	50.28	83.3(23.0)
P65-cpSSR_29	16.10	52.00	1.2(0.1)

P66-cpSSR_30	15.45	54.00	2.6(0.2)
P67-cpSSR_31	18.12	54.18	79.3(18.7)
P69-cpSSR_32	23.25	53.52	40.3(3.7)
P70-cpSSR_33	23.40	52.45	17.1(1.0)
P71-cpSSR_34	25.10	55.32	48.1(5.3)
P72-cpSSR_35	22.10	57.20	66.2(10.6)
P73-cpSSR_36	24.90	57.20	84.0(23.9)
P74-cpSSR_37	28.30	57.00	89.3(32.9)
P75-cpSSR_38	28.00	53.50	43.6(4.3)
P76-cpSSR_39	32.00	56.30	85.6(26.1)
P77-cpSSR_40	33.82	61.18	98.3(63.8)
P78-cpSSR_41	37.30	56.20	96.5(52.8)
P82-cpSSR_42	9.21	54.54	24.5(1.6)
P83-cpSSR_43	10.39	59.58	74.0(14.5)
P84-cpSSR_44	11.53	60.38	91.4(38.2)
P85-cpSSR_45	10.35	64.00	5.2(0.3)
P86-cpSSR_46	13.57	55.43	8.4(0.5)
P87-cpSSR_47	13.06	57.02	72.7(13.7)
P88-cpSSR_48	13.61	59.55	54.3(6.9)
P89-cpSSR_49	16.21	59.05	35.6(3.0)
P90-cpSSR_50	12.37	58.58	45.2(4.6)
P91-cpSSR_51	14.10	62.36	82.6(22.2)
P92-cpSSR_52	16.40	63.01	82.7(22.3)
P93-cpSSR_53	19.67	63.50	72.6(13.6)
P94-cpSSR_54	24.03	60.14	71.3(12.9)
P95-cpSSR_55	27.15	61.14	95.6(49.3)
Acc-cpSSR_56	7.82	44.35	0.0(0.0)
Ade-cpSSR_57	7.56	46.49	88.1(30.3)
Ale-cpSSR_58	8.02	46.38	67.2(11.0)
Aye-cpSSR_59	7.64	46.19	45.0(4.6)
Bod-cpSSR_60	8.85	46.98	93.3(42.4)
Bon-cpSSR_61	9.74	46.37	2.7(0.2)
Brs-cpSSR_62	7.30	45.46	70.2(12.3)
Bla-cpSSR_63	23.57	42.34	3.0(0.2)
Blb-cpSSR_64	23.50	41.83	0.0(0.0)
Blc-cpSSR_65	24.45	41.43	0.0(0.0)
Cmp-cpSSR_66	11.27	44.18	0.2(0.0)
Con-cpSSR_67	9.94	46.79	2.3(0.2)
Gst-cpSSR_68	7.32	46.34	33.3(2.6)
Ha-cpSSR_69	16.53	47.66	25.6(1.7)
Hb-cpSSR_70	20.06	47.85	1.0(0.1)
Hen-cpSSR_71	12.02	57.70	42.9(4.1)
Hyy-cpSSR_72	24.68	62.00	82.5(22)
Leb-cpSSR_73	6.40	46.66	73.4(14.1)
Lek-cpSSR_74	12.87	59.92	77.8(17.3)
Mon-cpSSR_75	7.01	46.49	80.4(19.8)
Mgx-cpSSR_76	7.20	45.80	77.9(17.4)

Nor-cpSSR_77	15.55	56.77	32.8(2.6)
Ors-cpSSR_78	7.23	46.07	38.2(3.3)
Plb-cpSSR_79	23.18	52.91	20.5(1.3)
Pos-cpSSR_80	10.19	46.33	38.5(3.4)
Ra-cpSSR_81	25.77	46.66	54.0(6.8)
Ron-cpSSR_82	9.76	46.63	53.8(6.7)
Sax-cpSSR_83	7.80	46.63	86.7(27.8)
Sbe-cpSSR_84	9.18	46.54	11.6(0.6)
Sca-cpSSR_85	9.20	46.79	97.3(56.6)
Sim-cpSSR_86	8.00	46.24	61.1(8.8)
Sku-cpSSR_87	15.07	57.87	47.1(5.0)
Slb-cpSSR_88	13.80	46.34	3.8(0.2)
Sve-cpSSR_89	13.12	57.20	76.4(16.2)
Tro-cpSSR_90	15.52	56.70	27.4(1.9)
Vld-cpSSR_91	7.70	44.25	18.4(1.1)
Vls-cpSSR_92	8.30	46.07	0.1(0.0)
SIJM-IZO_1	13.99	46.33	88.6(31.4)
SIJS-IZO_2	13.99	46.33	88.6(31.4)
SI-TB-IZO_3	14.12	46.33	6.5(0.4)
SI-LIP-IZO_4	13.93	46.37	87.1(28.5)
SI-BG-IZO_5	14.07	46.28	87.8(29.7)
SI-LED-IZO_6	14.11	46.25	90.2(35.1)
SI-LP-IZO_7	14.17	46.31	83.3(23.0)
SI-VV-IZO_8	14.44	46.40	89.3(32.9)
SI-PP-IZO_9	14.36	46.42	88.0(30.1)
SI-RK-IZO_10	14.67	46.37	92.2(40.0)
SI-DP-IZO_11	14.71	46.32	37.6(3.2)
SI-VP-IZO_12	14.66	46.29	79.9(19.3)
SI-CJ-IZO_13	15.44	46.44	93.7(43.4)
SI-TZ-IZO_14	15.41	46.47	93.1(42.0)
SI-KOM-IZO_15	15.32	46.44	92.2(40.0)
SI-SH-IZO_16	15.29	46.46	93.8(43.6)
SI-SMR-IZO_17	13.82	45.95	83.7(23.5)
SI-SD-IZO_18	13.87	45.97	87.7(29.5)
SI-GD-IZO_19	14.42	45.57	89.9(34.4)
SI-TR-IZO_20	14.42	45.54	90.0(34.6)
SI-VPAD-IZO_21	14.44	45.54	89.9(34.4)
SI-GRC-IZO_22	14.45	45.57	90.2(35.1)
SI-GS-IZO_23	14.74	45.58	88.4(31.0)
HR-SD-IZO_24	14.33	45.31	-
HR-MP-IZO_25	14.89	45.27	80.1(19.5)
HR-STI-IZO_26	15.06	44.68	75.1(15.3)
HR-VL-IZO_27	15.09	44.77	77.4(17.0)
BIH-GRM-IZO_28	16.62	44.66	35.4(2.9)
BIH-PRE-IZO_29	16.56	44.16	78.2(17.7)
BIH-BUS-IZO_30	16.82	44.03	67.3(11.0)
BIH-RAS-IZO_31	17.29	44.06	80.9(20.3)

BIH-VLA-IZO_32	17.45	44.32	66.8(10.8)
BIH-BIS-IZO_33	17.68	43.99	71.7(13.1)
BIH-IGMA-IZO_34	18.27	43.74	79.2(18.6)
BIH-IGMB-IZO_35	18.27	43.75	79.5(18.9)
BIH-ZEL-IZO_36	18.62	43.66	75(15.2)
BIH-BIJ-IZO_37	18.49	44.08	50.5(5.9)
BIH-TIB-IZO_38	18.34	44.34	0.2(0.0)
BIH-ROM-IZO_39	18.66	43.90	78.4(17.8)
BIH-HK-IZO_40	18.89	44.04	63.1(9.4)
D-OBA-IZO_41	10.97	47.60	82.4(21.9)
D-OBB-IZO_42	10.98	47.61	82.4(21.9)
D-OBC-IZO_43	10.97	47.61	82.4(21.9)
D-KOA-IZO_44	12.96	47.80	25.9(1.8)
D-KOB-IZO_45	12.96	47.81	25.9(1.8)
F-BOA-IZO_47	6.27	45.11	7.6(0.4)
F-BOB-IZO_48	6.30	45.11	3.8(0.2)
F-NEB-IZO_49	6.64	45.01	63.4(9.5)
F-NEC-IZO_50	6.63	45.01	55.9(7.3)
F-SFTA-IZO_51	6.81	45.61	80.2(19.6)
F-SFTB-IZO_52	6.79	45.60	80.6(20.0)
F-SFTC-IZO_53	6.81	45.59	43.5(4.3)
APS-1-IZO_54	22.77	46.60	79.0(18.4)
APS-2-IZO_55	22.76	46.58	80.6(20.0)
APS-3-IZO_56	22.75	46.61	86.0(26.7)
PRG-1-IZO_57	23.64	45.41	47.7(5.2)
PRG-2-IZO_58	23.62	45.41	67.8(11.2)
PST-1-IZO_59	25.56	45.57	70.8(12.6)
PST-2-IZO_60	25.54	45.61	43.2(4.2)
MMS-1-IZO_61	24.62	47.76	60.9(8.7)
MMS-2-IZO_62	24.62	47.75	60.9(8.7)
NMR-IZO_63	26.35	46.15	64.0(9.7)
PRS-IZO_64	22.52	45.72	63.8(9.7)