Departamento de Botánica

UNIVERSIDAD DE GRANADA

Tesis Doctoral

Ph.D. Thesis

Diversidad genética y modelo filogeográfico de cuatro especies de helechos relictos del Terciario: *Vandenboschia speciosa* (Willd.) G. Kunkel, *Diplazium caudatum* (Cav.) Jermy, *Pteris incompleta* Cav. y *Culcita macrocarpa* C. Presl.

Genetic diversity and phylogeographic pattern of four Tertiary-relict fern species: *Vandenboschia speciosa* (Willd.) G. Kunkel, *Diplazium caudatum* (Cav.) Jermy, *Pteris incompleta* Cav. and *Culcita macrocarpa* C. Presl.

Samira Ben-Menni Schuler

Granada, 2019

Programa de Doctorado Biología Fundamental y de Sistemas

Editor: Universidad de Granada. Tesis Doctorales
Autor: Samira Ben-Menni Schuler
ISBN: 978-84-1306-402-4
URI: http://hdl.handle.net/10481/58616

Contents

Summary 1

Resumen 5

Introducción (General introduction; in Spanish)

- La Geoflora Paleotropical 9
- El uso de marcadores moleculares para investigar procesos históricos y contemporáneos en poblaciones vegetales 11
- Especies objeto de estudio 15
- Referencias 20

Chapters

Chapter 1. Genetic diversity and population history of the Killarney fern, *Vandenboschia speciosa* (Hymenophyllaceae), at its southern distribution limit in continental Europe. 25

Chapter 2. Phylogeographical analyses of a relict fern of palaeotropical geoflora (*Vandenboschia speciosa*): distribution and diversity model in relation to the geological and climate events of the late Miocene and early Pliocene. 51

Chapter 3. Genetic diversity and phylogeography of the relict fern *Culcita macrocarpa*: influence of clonality and reproduction system on genetic structure. 109

Chapter 4. Population variability and differentiation and phylogeographic history in the endemic Ibero-Macaronesian fern *Diplazium caudatum* (Athyriaceae). 153

Chapter 5. Development of polymorphic microsatellite markers for the Iberian-Macaronesian endemic Laurisilva Brake (*Pteris incompleta*, Pteridaceae). 193

Discusión (General discussion; in Spanish) 201

Conclusions 207

Conclusiones 209

Summary

One of the main challenges of the evolutionary biology is to know the causes and processes that at level of species have caused the different patterns of diversity, distribution, and diversification of populations and gene lineages. The development and improvement of population genetics and phylogeographic analyses in relation to the advance in molecular biology techniques over the last decades has allowed these challenges to be overcome. However, there is still a large number of groups of organisms in which the phylogeographic and diversity pattern of many of its species has not been studied in depth, such as ferns and lycophytes. Fern and lycophyte species belong to ancestral lineages between the vascular plants, being the ferns the second most abundant group after angiosperms, with about 10,578 species. This doctoral thesis is developed with the objective of elucidating the patterns of genetic diversity and phylogeography of four species of Tertiary relict ferns: *Culcita macrocarpa*, *Diplazium caudatum*, *Pteris incompleta*, and *Vandenboschia speciosa*. These four ferns have a common history, having hypothetized their past belonging to the Paleotropical geoflora that populated the middle and lower latitudes of Eurasia and the western America from the late Cretaceous to the Upper Miocene. Currently, they share their distribution area and ecological niche, living toghether in most locations, in disjunct areas considered glacial refugia. Various molecular markers, of nuclear and/or chloroplast DNA origin, have been used to accomplish this objective.

In Chapter 1 we investigate the genetic diversity and structure of *Vandenboschia speciosa*, at its southernmost distribution limit in continental Europe (S Iberian Peninsula: Andalusia, Spain). We analysed sporophyte populations using ten microsatellite loci. The results indicated moderate genetic diversity at the regional scale but low diversity at the population level and a significant genetic differentiation between populations. Population history in Andalusia is characterized by a migration-drift equilibrium with historical dispersals as the main factors influencing population structure.

In Chapter 2 we reconstruct the evolutionary history of *Vandenboschia speciosa*, hypothesised to be of Tertiary Palaeotropical geoflora with a peculiar perennial gametophyte. We sequenced 40 populations across the species range in one plastid

region and two paralogues of the nuclear *gapCp* gene and conducted time-calibrated phylogenetic, phylogeographical, and species-distribution modelling analyses. *Vandenboschia speciosa* had a Tertiary origin. Late Miocene aridification possibly caused the long persistence in independent refugia on Eurosiberian Atlantic and Mediterranean coasts, with the independent evolution of gene pools resulting in two evolutionary units. Our results support the idea that the geological and climate events of the late Miocene/early Pliocene shifted Tertiary fern distribution patterns in Europe.

In Chapter 3 we infer the phylogeographic history of *C. macrocarpa*, and evaluate the importance and effect of breeding system and clonality on the genetic variability of the species. We provided genetic data from eight microsatellite nuclear loci and one plastid DNA region for 17 populations covering the distribution range of the species, and species-distribution modelling data. The global phylogeography of *C. macrocarpa* is characterized by the differentiation of populations in two main groups, which coincide with the geographical distribution of the species in the North of the Iberian Peninsula and in the Macaronesian archipelagos and southern Iberian Peninsula. The differentiation of the gene lineages that characterize each group began during que Early Miocene (19.5 Ma). Within each group of populations we have detected, based on the levels of diversity, prevalence of clonality and interpopulation differentiation, relict populations (in Azores and several of the Cantabrian Cornice) and populations of recent origin. This population structure responds to a dynamic of colonization in which recent populations have originated from one or a few genets, from relictual populations, which have been established by intragametophytic selfing and vegetatively expanded. Current Andalusian populations have appeared by multiple colonizing events from Azores and the Cantabrian Cornice.

In Chapter 4 we characterize the genetic diversity and the phylogeographic pattern of *Diplazium caudatum* throughout its distribution area. We analysed 12 populations using eight microsatellite loci, sequence one plastid region, and carry out species-distribution modelling analyses. We confirmed the Tertiary origin of these species. It seems that *D. caudatum* took refuge in the Macaronesian archipelagos when the coldest and driest climate began during the Pliocene-Pleistocene transition, and subsequently recolonized the continent, being present only in Andalusia. The model that best explains the pattern

of distribution and diversity observed in this species is long-distance dispersion with restriction of gene flow and genetic drift.

In Chapter 5 we characterize eight microsatellite loci in *Pteris incompleta*, enabling future studies on the population genetic of this Iberian-Macaronesian species. The eight primer sets were tested on a total of 47 individuals from one Azorean and one Madeiran population, resulting in medium-high levels of diversity.

Resumen

Uno de los principales desafíos de la biología evolutiva es conocer las causas y procesos que a nivel de especies han causado los diferentes patrones de diversidad, distribución y diversificación de poblaciones y linajes genéticos. El desarrollo y la mejora de la genética de poblaciones y los análisis filogeográficos en relación con el avance de las técnicas de biología molecular en las últimas décadas ha permitido superar estos desafíos. Sin embargo, todavía hay una gran cantidad de grupos de organismos en los que el patrón filogeográfico y de diversidad de muchas de sus especies no se ha estudiado en profundidad, como los helechos y los licófitos. Las especies de helechos y licófitos pertenecen a linajes ancestrales entre las plantas vasculares, siendo los helechos el segundo grupo más abundante después de las angiospermas, con aproximadamente 10.578 especies. Esta tesis doctoral se desarrolla con el objetivo de dilucidar los patrones de diversidad genética y filogeografía de cuatro especies de helechos relictos terciarios: *Culcita macrocarpa*, *Diplazium caudatum*, *Pteris incompleta* y *Vandenboschia speciosa*. Estos cuatro helechos tienen una historia común, habiendo hipotetizado su pasado perteneciente a la geoflora paleotropical que pobló las latitudes medias y bajas de Eurasia y el oeste de América desde el Cretácico tardío hasta el Mioceno superior. Actualmente, comparten su área de distribución y su nicho ecológico, viviendo juntos en la mayoría de los lugares, en áreas disyuntas consideradas refugios glaciales. Se han utilizado diversos marcadores moleculares, de ADN de origen nuclear y/o cloroplastidial, para lograr este objetivo.

En el Capítulo 1 investigamos la diversidad genética y la estructura de *Vandenboschia speciosa*, en su límite de distribución más meridional en Europa continental (S Península Ibérica: Andalucía, España). Analizamos las poblaciones de esporofitos utilizando diez loci microsatélites. Los resultados indicaron una diversidad genética moderada a escala regional pero una diversidad baja a nivel de la población y una diferenciación genética significativa entre las poblaciones. La historia poblacional de *V. speciosa* en Andalucía se caracteriza por un equilibrio migración-deriva génica con dispersiones históricas como los principales factores que influyen en la estructura de las poblaciones.

En el Capítulo 2 reconstruimos la historia evolutiva de *Vandenboschia speciosa*, que se supone que es de origen Terciario y que perteneció a la Geoflora Paleotropical terciaria con un gametofito perenne peculiar. Muestreamos 40 poblaciones en todo el rango de distribucón de la especie y secuenciamos una región cloroplastidial y dos parálogos del gen gapCp nuclear y realizamos análisis filogenéticos calibrados en el tiempo, filogeográficos y modelos de distribución de especies. *Vandenboschia speciosa* tuvo un origen terciario. La aridificación del Mioceno tardío posiblemente causó la larga persistencia en refugios independientes de las costas atlánticas y mediterráneas de Eurosiberia, con la evolución independiente de los pools génicos que dieron como resultado dos unidades evolutivas. Nuestros resultados apoyan la idea de que los eventos geológicos y climáticos del Mioceno tardío / Plioceno temprano cambiaron los patrones de distribución de helechos terciarios en Europa.

En el Capítulo 3 inferimos la historia filogeográfica de *C. macrocarpa*, y evaluamos la importancia y el efecto del sistema de reproducción y la clonalidad sobre la variabilidad genética de la especie. Proporcionamos datos genéticos de ocho loci microsatélites nucleares y una región de ADN cloroplastidial para 17 poblaciones que cubren el rango de distribución de la especie, y datos de modelado de distribución de especies. La filogeografía global de *C. macrocarpa* se caracteriza por la diferenciación de las poblaciones en dos grupos principales, que coincide con la distribución geográfica de las especies en el norte de la Península Ibérica y en los archipiélagos macaronesios y el sur de la Península Ibérica. La diferenciación de los linajes génicos que caracterizan a cada grupo comenzó durante el Mioceno Temprano (19.5 Ma). Dentro de cada grupo de poblaciones hemos detectado, en base a los niveles de diversidad, prevalencia de clonalidad y diferenciación interpoblacional, poblaciones relictas (en Azores y varias de la Cornisa Cantábrica) y poblaciones de origen reciente. Esta estructura poblacional responde a una dinámica de colonización en la que las poblaciones recientes se han originado a partir de uno o unos pocos genets, de poblaciones relictas, que se han establecido por autofecundación intragametofítica y expandido vegetativamente. Las poblaciones andaluzas actuales han aparecido por múltiples eventos de colonización de Azores y la Cornisa Cantábrica.

En el Capítulo 4 caracterizamos la diversidad genética y el patrón filogeográfico de *Diplazium caudatum* en toda su área de distribución. Analizamos 12 poblaciones

utilizando ocho loci de microsatélites, secuenciamos una región cloroplastidial y realizamos análisis de modelos de distribución de especies. Confirmamos el origen terciario de esta especie. Parece que *D. caudatum* se refugió en los archipiélagos macaronésicos cuando comenzó el clima más frío y seco durante la transición del Plioceno-Pleistoceno, y posteriormente recolonizó el continente, estando presente solo en Andalucía. El modelo que mejor explica el patrón de distribución y diversidad observado en esta especie es la dispersión a larga distancia con restricción del flujo genético y la deriva genética.

En el Capítulo 5 caracterizamos ocho loci microsatélites en *Pteris incompleta*, lo que permitirá futuros estudios sobre la genética de poblaciones de esta especie ibérico-macaronésica. Los ocho conjuntos de cebadores se probaron en un total de 47 individuos de una población de Azores y una de Madeira, lo que resultó en niveles de diversidad medio-altos.

Introducción

1. La Geoflora Paleotropical

El periodo Terciario se desarrolló a partir de un evento de extinción masiva que tuvo lugar hace 65 m.a. (millones de años). El Terciario se encuentra dividido en dos grandes sistemas: el Paleógeno (65 – 23.3 m.a.) y el Neógeno (23.3 – 1.64 m.a.). El Paleógeno puede dividirse en tres series: Paleoceno, Eoceno y Oligoceno; dividiéndose el Neógeno en dos: Mioceno y Plioceno. Posteriormente se encuentra el Cuaternario dividido en el Pleistoceno (2.58 m.a. – 12.000 años) y el Holoceno (12.000 – actualidad). Este evento afectó a muchos animales, mientras que las plantas experimentaron la extinción de solo unos pocos taxones.

El Reino Holártico durante gran parte del Terciario estuvo poblado por dos cinturones de vegetación, denominados Geofloras (Mai, 1991). Una de ellas es la Geoflora Arctoterciaria que presenta un clima templado cálido y se distribuye en lo que actualmente es el noreste de Europa, Siberia, Este de Rusia, parte de Japón, Corea, Norteamérica y regiones árticas. Las plantas características de esta geoflora se pueden encontrar en los bosques mesófilos caducifolios y de coníferas del Hemisferio Norte (Barrón, 2003).

En las regiones al sur de la zona Ártica, en latitudes medias y bajas del oeste de Norteamérica y Eurasia (Estados Unidos, centro y sur de Europa, Asia Menor, la mayor parte de Kazakhstán, Asia Central, China y parte Japón), se desarrolló hasta el Mioceno Superior una geoflora caracterizada por bosques con requerimientos paratropicales y subtropicales lluviosos, laurisilvas y bosques templados integrados por taxones perennifolios, que soportaban climas monzónicos. El estrato herbáceo de estos bosques estaba poblado fundamentalmente por helechos (Barrón, 2003). Ésta es la denominada Geoflora Paleotropical.

El área donde se desarrolló la Geoflora Paleotropical abarcó todo el Hemisferio Norte a lo largo de las costas del Thethys (antiguo mar que comunicaba el océano Atlántico y el Índico). Parece ser que en Europa se originó y desarrolló una flora de tipo lauroide exclusiva ya que la región europea estuvo aislada durante el Paleoceno del resto de las masas continentales euroasiáticas y americanas por el mar de Tethys y los océanos Ártico y Atlántico. Estos bosques lauroides siempreverdes estuvieron

presentes en Eurasia casi todo el Paleógeno y el Neógeno temprano y tenían requerimientos climáticos de elevada humedad y calor (Mai, 1989, 1991).

Durante el Mioceno Medio (11-16 m.a.) y el Mioceno Superior (11-3.4 m.a.) comienza una mayor continentalización del clima, debido al enfriamiento de las temperaturas y el aumento de la aridez, y se produce un cambio climático en Europa que pasa a ser de subtropical-monzónico a templado. Esto conllevó a un retroceso hacia el sur y la extinción de muchos taxones lauroides. Posteriormente, debido a la crisis de salinidad que provocó el cierre del Mediterráneo, se desarrolló un clima árido que tuvo como consecuencia la formación de zonas esteparias. El último periodo del Terciario es el Plioceno (5.3-1.64 m.a.) en el cual se instalaron las condiciones climáticas semejantes a las actuales en la cuenca mediterránea y como consecuencia, se produjo la extinción de distintos taxones típicos del Terciario, como la familia Taxodiaceae (Barrón 2003). Es por ello que los bosques perimediterráneos de tipo tropical han quedado acantonados en aquellos lugares que han ofrecido condiciones más favorables para su supervivencia, como sería la región Macaronésica (Azores, Islas Canarias y Madeira) donde llegaron probablemente entre el Mioceno y el Plioceno (Axelrod, 1990), mientras que el carácter más árido del sur de Europa determinó el desarrollo de la actual vegetación esclerófila mediterránea (Ferreras & Arocena, 1987). Más tarde, además, las glaciaciones del Pleistoceno redujeron drásticamente las áreas habitables para la permanencia de taxones como por ejemplo *Laurus*, provocando la extinción de poblaciones relictas, particularmente en latitudes más al norte (Europa central y sudeste). Durante estos periodos glaciales, algunos de estos taxones pudieron haber persistido en múltiples pero pequeños y disyuntos refugios distribuidos por los archipiélagos macaronésicos y la cuenca mediterránea, como la Península Ibérica, el norte de África, el sur de Italia y la región Balcano-Egea como principales refugios glaciales. El rango actual de distribución de estas especies está constreñido por el clima: el frio impide la expansión hacia el norte y la sequía pone el límite hacia sur, viviendo en zonas costeras muy húmedas y templadas (Rodriguez-Sánchez & Arroyo, 2008). Sin embargo, se ha comprobado la existencia de refugios durante las glaciaciones a latitudes más elevadas (Stewart & Lister, 2001; Provan & Bennett, 2008) pudiendo expandir estas especies posteriormente sus poblaciones (Taberlet et al., 1998; Hewitt, 1999; Stewart & Lister, 2001).

Se consideran especies relictas del Terciario las localizadas en auténticos refugios que conservan condiciones microclimáticas similares a las de aquella época; y que muestran un declive poblacional importante ya sea por el cambio en estas condiciones microclimáticas como por la acción del hombre (*cf.* Delgado Vázquez & Plaza Arregui, 2006).

Las especies de helechos pertenecen a linajes ancestrales entre las plantas vasculares (Schneider et al., 2004). En Europa este grupo de plantas estuvo ampliamente representado durante el Terciario, sin embargo los factores geológico-climáticos acaecidos durante el Plioceno y Pleistoceno conllevaron la extinción de muchos de estos linajes debido a la disminución de la temperatura y el aumento de la aridez que culminaron en las glaciaciones del Cuaternario (Barrón, 2003; Rodríguez-Sánchez y Arroyo, 2008; Postigo Mijarra et al., 2009; Fernández-Palacios et al., 2011). Esto ha provocado que muchas especies que sobrevivieron a estos eventos quedaran relegadas a zonas disyuntas climáticamente favorables de la franja atlántica europea y especialmente de la Península Ibérica e islas Macaronésicas. Como consecuencia de la reducción y fragmentación de sus poblaciones muchas especies de helechos europeos se encuentran catalogadas como especies amenazadas.

2. El uso de marcadores moleculares para investigar procesos históricos y contemporáneos en poblaciones vegetales.

Los marcadores moleculares son fragmentos de ADN que pueden ser identificados y cuantificados dentro del genoma de forma relativamente sencilla. Por ello, se han utilizado en numerosas ocasiones como base para la identificación de especies, elaboración de árboles génicos, estudios de biología evolutiva o análisis de diversidad genética (Avise, 1994; Cheng & Crittenden, 1994). Para el desarrollo de este trabajo se han utilizado como marcadores moleculares secuencias microsatélites del ADN nuclear y secuencias nucleotídicas del ADN cloroplastidial y nuclear.

Microsatélites

La diversidad genética, se entiende como una medida de los diferentes alelos provistos por un gen, los cuales coexisten en los diferentes individuos y en las diferentes poblaciones. En consecuencia, un número elevado de alelos presente en los diferentes individuos de una determinada población, es significativo de una diversidad genética elevada. En general, y

para las especies amenazadas en particular, la diversidad genética es crucial para mantener sus expectativas evolutivas y representa la base que asegura su capacidad reproductiva, su resistencia a enfermedades y su capacidad de adaptación a los cambios ambientales. Una población sin variabilidad genética no puede evolucionar ni adaptarse a los cambios producidos en el ambiente, lo cual aumenta su riesgo de extinción. Documentar la diversidad genética y su distribución y el mantenimiento de esa diversidad genética es un factor importante para la supervivencia y gestión de la conservación de las especies amenazadas. Para analizar y obtener los niveles de diversidad genética y cómo se estructura esa diversidad en las poblaciones de las cuatro especies de helechos estudiadas se han utilizado en el presente trabajo marcadores moleculares microsatélites. Los microsatélites son secuencias de ADN no codificante constituidas a partir de motivos (2-10 pares de bases de longitud) repetidos en tándem (> 4 repeticiones) que se encuentran a lo largo de todo el genoma de los eucariotas (Chambers & MacAvoy, 2000). Para cada individuo, la variación en el número de repeticiones para un locus microsatélite crea diferentes alelos. Las secuencias microsatélites han demostrado su utilidad en numerosos estudios de genética poblacional, habiéndose convertido en el principal marcador de elección en este tipo estudios (Ellegren, 2004; Fernandez-Silva, 2013; Jarne & Lagoda, 1996; Li et al., 2002; Selkoe & Toonen, 2006; Zhang & Hewitt, 2003). Esto se debe a sus características: i) poseen una alta tasa de mutación debida a fallos durante el proceso de replicación del ADN, por un mal apareamiento entre unidades repetidas (Mohan et al., 1997; Oliveira et al., 2006) lo que los convierte en muy polimórficos, es decir, presentan una elevada cantidad de variantes alélicas, lo que los hace muy convenientes para definir un único genotipo multilocus, ii) son neutros, no se ven afectados por la selección natural, iii) son codominantes, por lo que permiten discriminar entre homocigotos y heterocigotos, iv) son fáciles de aislar y amplificar por PCR, ya que poseen menos de 300 pb, y v) son altamente reproducibles (Brown, 2007; Jarne & Lagoda, 1996; Oliveira et al., 2006). Estas características convierten a los microsatélites en herramientas muy útiles para realizar estimaciones de la diversidad genética y de la consanguinidad existente en poblaciones.

Secuencias cloroplastidiales

Otro factor importante en la conservación de especies amenazadas es identificar los factores históricos responsables de la actual distribución de las especies, y cómo es la dinámica poblacional. El análisis espacial de los linajes génicos, es decir, la filogeografia permite tal

identificación. La filogeografía aplica el análisis de genealogías génicas al estudio de la evolución de las poblaciones y permite sacar conclusiones con respecto a las secuencias de colonización, diversificación, aislamiento y extinción de los linajes génicos en determinadas áreas. Además, el estudio comparado de los patrones filogeográficos de varias especies codistribuidas contribuye a plantear hipótesis sobre posibles eventos comunes de vicarianza o dispersión y a identificar las causas geológicas, ecológicas o etológicas que pudieran haber influido en ellos (*cf.* Lanteri & Confalonieri, 2003). En los estudios filogeográficos es necesario utilizar marcadores moleculares para datar los linajes génicos y reconstruir las relaciones filogenéticas entre ellos. Se usan marcadores cloroplastidiales debido a la alta tasa de mutación, lo que proporciona buena resolución en estudios a nivel intraespecífico, y a que tienen herencia uniparental, por lo que no hay recombinación sexual y nos permite reconstruir las relaciones filogenéticas de manera linear y poder establecer genealogías génicas (Domínguez–Domínguez & Vázquez, 2009). Si se trabaja con muestras de genes nucleares provenientes de individuos entre los cuales existe o existió alguna vez flujo génico, es altamente probable que haya habido recombinación genética durante el apareamiento de cromosomas homólogos en la meiosis. Este proceso puede llegar a desvirtuar las verdaderas genealogías de las secuencias. Las secuencias utilizadas como marcadores filogeográficos deberán ser variables intraespecíficamente de forma que ofrezcan resolución suficiente para realizar el análisis.

Evolutivamente hablando, estos marcadores moleculares (microsatélites y secuencias cloroplastidiales) pueden darnos información acerca de las fuerzas evolutivas que han podido actuar en las poblaciones de determinadas especies, como pueden ser las mutaciones, deriva génica y migración (Lowe et al., 2004). Las mutaciones son cambios en la secuencia del DNA provocando la creación de variantes alélicas e incrementando la diversidad genética. La deriva génica son fluctuaciones al azar de las frecuencias alélicas de generación en generación y provoca que las poblaciones se vuelvan diferentes genéticamente entre ellas simplemente por azar. Cuanta más pequeña es una población, más poderosa es la deriva. Esto está contrarrestado por la migración, que mantiene la conectividad entre individuos y poblaciones. Sin embargo, si el flujo génico entre poblaciones está restringido, éstas se diferenciarán y acabarán siendo genéticamente diferentes. Hay cierta aceptación general respecto a que la reducción del tamaño poblacional efectivo da lugar a una reducción de la diversidad genética (Widmer & Lexer, 2001), por ejemplo dispersiones a nuevos hábitats

podrían dar lugar al llamado efecto fundador ya que esas nuevas poblaciones consisten en uno o pocos individuos que representan una muestra muy pequeña de la diversidad genética de la población de origen. Reducciones en el tamaño poblacional también pueden ocurrir por la destrucción de hábitat y en el caso de que la población recobrara su tamaño poblacional, la diversidad genética es una muestra pequeña de la diversidad original. Este proceso es llamado cuello de botella.

Además, los resultados que podemos obtener de estos marcadores moleculares también podrían ayudarnos a dilucidar otras posibles causas, además de las fuerzas evolutivas, que hayan podido afectar a ese patrón de diversidad y modelo filogeografico, como podrían ser eventos geológico-climáticos pasados, como por ejemplo, las glaciaciones del Cuaternario. La mayoría de estudios filogeográficos en plantas, analizan los efectos de las glaciaciones del Pleistoceno sobre la distribución de las poblaciones pero no sucesos más antiguos en taxones como los encontrados en los bosques lauroides. En helechos son pocos los estudios que se han llevado a cabo en Europa continental y Macaronesia (e.g. Trewick et al., 2002; Hunt et al., 2009; Jiménez et al., 2010; Bystriakova et al., 2014, Maccagni et al., 2017) y ninguno se ha realizado para reconstruir la historia biogeográfica de una posible especie de helecho del Terciario. El modelo de distribución de estos helechos hace de ellos especies de interés para explorar las respuestas de las plantas en europa a los cambios ambientales y la huella genética dejada por estos cambios.

Las cuatro especies estudiadas en esta Tesis Doctoral son: *Vandenboschia speciosa*, *Culcita macroarpa*, *Diplazium caudatum* y *Pteris incompleta*. Se han elegido estas especies en concreto porque tienen en común que se consideran relictas del Terciario y coinciden en área de distribución y en nicho ecológico, de forma que conviven en la mayoría de las localidades. La discontinuidad entre las poblaciones de estas especies, la vulnerabilidad de los hábitats en los que se encuentran y el escaso número de individuos de sus poblaciones representan un elevado riesgo de erosión genética. Las cuatro están catalogadas como especies amenazadas en distintos libros o catálogos, en los ámbitos europeo y nacional. Esta situación hace de ellos candidatos prioritarios para realizar estudios que permitan establecer cómo de depauperada se encuentra su diversidad genética y así poder incluir esta información en los planes de gestión de los mismos. Además, el poder analizar y contrastar los patrones de diversidad y filogeográficos entre especies del mismo grupo y codistribuidos

geográficamente permite establecer unos patrones comunes de posibles causas y eventos históricos que pudieron dar lugar a la situación actual de las poblaciones de estas especies.

3. Especies objeto de estudio

En los helechos, el ciclo de vida implica fases haploides (gametofito) y diploides (esporofito) independientes, y la fertilización ocurre en el gametofito originado de las esporas. Es por ello, que la fertilización siempre es un proceso post-dispersivo y por lo tanto la probabilidad de colonización va a depender del sistema de reproducción de la especie. Es menos probable para helechos con reproducción cruzada o reproducción inter-gametofítica que para aquellos con reproducción intra-gametofítica ya que son capaces de establecerse en una nueva población con una sola espora (Suter et al., 2000; de Groot et al., 2012). Sin embargo, muchos helechos tienen la capacidad de reproducirse vegetativamente en alguna o ambas generaciones, lo que afecta al proceso de colonización y por lo tanto a la estructura genética poblacional (Johns & Edwards, 1991; Rumsey et al., 1999).

Además, la dinámica de colonización de estos helechos terciarios depende de la disponibilidad de hábitats adecuados con humedad y temperaturas altas con una fuerte influencia oceánica. Sin embargo, esos hábitats suelen estar disyuntamente distribuidos y como consecuencia, limita la dispersión y conectividad de las poblaciones incluso a pesar de que los helechos tengan una elevada capacidad dispersiva (e.g. Bystriakova et al., 2014; Hunt et al., 2009; Jiménez et al., 2010; Maccagni et al., 2017; Shepherd et al., 2007; Trewick et., al 2002; Wang et al., 2011)

1- Vandenboschia speciosa

Vandenboschia speciosa pertenece a la familia Hymenophyllaceae. Es un endemismo Macaronésico-Europeo raro, único representante en la zona de un género de distribución principalmente tropical (Rumsey et al., 1998) que se considera relicto del Terciario (Barrón & Peyrot, 2006). El esporofito es rizomatoso, capaz de progaparse por fragmentación de su rizoma. Las frondes están insertadas a lo largo del rizoma y las láminas son delgadas, translúcidas, constituidas por una sola capa de células y verde oscura. La especie es tetraploide 2n=144. El gametofito, muy diferente al típico protalo con forma corazonada, se caracteriza por ser epigeo y angostamente taloso o filamentoso (hasta tal punto que suele confundirese con el protonema de un briofito), además de productor de yemas geminíferas.

La especie se caracteriza porque las dos fases de su ciclo de vida son perennes y capaces de reproducirse mediante propagación vegetativa (Rumsey et al., 1999). Es por ello, que existen poblaciones con ambas generaciones o independientes a lo largo del rango de distribución de la especie (Rumsey et al., 1998). Mientras que el esporofito está adaptado a crecer en zonas con escasa incidencia de luz y con constante humedad, el gametofito puede vivir en un rango más amplio de hábitats, inluyendo aquellos menos húmeros y más oscuros. Tales sitios pueden proveer un microclima y un ambiente estable para la supervivencia durante largos periodos de tiempo de los gametofitos independientes, fuera del rango de distribución del esporofito (Rumsey et al., 1999). Las poblaciones con esporofitos y gametofitos están restringidas a la costa atlántica europea y las islas macaronésicas (Azores, Canarias y Madeira), en lugares considerados refugios (del terciario y las glaciaciones) llevando a cabo ciclos de vida normales (Rumsey et al., 2005), mientras que poblaciones gametofíticamente independientes pueden encontrarse en el centro y norte de Europa (Rumsey et al., 2005) donde el esporofito desaparece. Está considerada como una de las especies de helechos más vulnerables de Europa, amenazada por la destrucción del hábitat y por la recolección excesiva, estando recogida en el Anexo I de la Convención de Berna y en el anexo II de la Directiva Hábitats (Anonymous, 1979, 1992). A nivel nacional está considerada como vulnerable en la Lista Roja de la Flora Vascular Española 2000. Hasta el momento se ha llevado a cabo un solo análisis sobre la filogeografía de la especie realizado por Rumsey et al. (1996), según el cual se encontraron dos haplotipos distintos usando PCR-RFLPs, estando distribuido uno de los haplotipos en el sur del rango de distribución (Península Ibérica, Italia, Madeira y las Islas Canarias) y el otro en el norte (Bretaña francesa e Irlanda) sugiriendo el retroceso y expansión de las poblaciones durante las glaciaciones como posibles causas de este resultado. Rumsey et al. (1998, 1999, 2005) realizaron varios análisis de genética poblacional utilizando isoenzimas en poblaciones del centro de Europa (Los Vosges y Alemania), Escocia, Andalucía e Italia. Con los resultados que obtuvieron, estos autores interpretaron que las poblaciones centroeuropeas y escocesas podrían ser fruto de recolonizaciones post-glaciales y las otras áreas de refugio. Como se apuntó en *C. macrocarpa*, los resultados obtenidos mediante los análisis con isoenzimas hay que tomarlos con cautela debido a la baja diversidad que suelen presentar estos marcadores moleculares.

2- Culcita macrocarpa

Culcita macrocarpa es una especie de helecho de la familia Culcitaceae que puede llegar a adquirir gran tamaño, con frondes de hasta 2,5 metros de longitud y limbo verde oscuro brillante y triangular, rizoma rastrero que puede llegar a alcanzar más de un metro de longitud y cubierto con páleas pluricelulares de color ferruginoso. Sus soros son marginales, incluidos dentro de unas estructuras ovoideas formadas por dos valvas (Delgado & Plaza, 2006). Es un hemicriptófito siempreverde, que se desarrolla y esporula durante todo el año. Sus yemas pueden permanecer de 3 a 4 meses inactivas, alcanzando el desarrollo máximo de la fronde en unas 3 semanas (Delgado & Plaza, 2006). Presenta un sistema de reproducción en el que la autogamia (incluyendo la automixis) se encuentra muy favorecida, así como un fuerte crecimiento clonal (Quintanilla et al., 2005). Se considera una especie diploide con 2n = 136 cromosomas (Manton et al., 1986). *C. macrocarpa*, es un helecho endémico de Portugal y España considerado relicto del Terciario; la mayor parte se encuentra en territorio portugués (Portugal continental, Azores y Madeira). En España se distribuye a lo largo de la Costa Cantábrica, el norte de Galicia, las Sierras de Algeciras (Cádiz) y la isla de Tenerife. Su hábitat ideal son los bosques de ribera en laderas orientadas al norte, en valles generalmente cerrados, a baja altitud y a una distancia relativamente cercana de la costa. Son bosques muy umbríos, con alta humedad atmosférica y sobre suelos muy húmedos y de carácter silíceo (Amigo & Norman, 1995). La especie se encuentra representada en diferentes figuras de protección y legislación a nivel internacional como son el Convenio de Berna y la Directiva Hábitats (Anonymous; 1986, 1992). A nivel nacional, *C. macrocarpa* aparece incluida en la Lista Roja de la Flora Vascular Española (Bañares et al., 2010) como en peligro. Existe un único trabajo que analiza la diversidad genética de *C. macrocarpa* y es el realizado por Quintanilla et al. (2007). Estos autores estudiaron la variabilidad genética en las poblaciones cantábricas de esta especie y detectaron una ausencia total de diversidad en las mismas, achacando este hecho a un fenómeno de deriva genética asociado a la reducción en el número de poblaciones durante la última era glaciar y a posteriores efectos fundadores a partir de pocas poblaciones. Estos resultados hay que tomarlos con cautela y merecen ser contrastados, pues las isoenzimas, marcador molecular utilizado en este estudio quizás no sean los mejores marcadores a emplear, dada la escasa resolución que ofrecen, cuando la diversidad genética que se espera es baja.

3- *Diplazium caudatum*

Diplazium caudatum es un helecho perteneciente a la familia Athyriaceae. Se caracteriza por poseer un rizoma rastrero de hasta 40 cm, con frondes desde 50 hasta 190 cm de largo y 25-50 cm de anchura. Dichas frondes son de color negro verdoso en la parte inferior y amarillo verdoso o plomizo en la superior, cubierto en la base de páleas castaño oscuras y con hojas persistentes. Presenta soros oblongos dispuestos a ambos lados de las venas, con indusio, y tiene esporas monoletas.

El desarrollo y esporulación de *D. caudatum* se produce a lo largo de todo el año. Se trata de una especie diploide, número cromosómico 2n = 82 (Blanca, 1999; Delgado y Plaza, 2006). El comportamiento fenológico está determinado por condiciones climáticas puntuales, siendo máximo el desarrollo cuando las condiciones de humedad y temperatura son favorables, produciendo nuevas frondes, yemas vegetativas y esporangios maduros (Delgado & Plaza, 2006; Moya et al., 2010). Es un endemismo Europeo-Macaronésico considerado relicto del Terciario que crece en sotobosques de Laurisilva y barrancos oscuros, sobre suelos muy húmedos, con gran humedad atmosférica permanente y temperaturas cálidas y constantes (Blanca, 1999; Delgado & Plaza, 2006; Moya et al., 2010). Actualmente se encuentra como relicto en áreas consideradas refugios durante los periodos glaciares del Cuaternario (Rumsey et al., 2005), en las Sierras de Algeciras (Cádiz), al sur de la Península Ibérica, y en las islas Macaronésicas, concretamente en las Islas Azores, Cabo Verde (solo descrita en la Isla de Santo Atâo), Canarias (Gran Canaria, La Gomera, La Palma y Tenerife) y Madeira (Blanca, 1999; Delgado & Plaza, 2006). En estas regiones el clima es de carácter subtropical y se dan las condiciones idóneas para que se desarrolle. El origen del género *Diplazium* se remonta al paleógeno hace más de 60 millones de años (Wei et al., 2015). El subgénero *Pseudallantodia* (Clarke), al que pertenece *D. caudatum*, es el linaje más basal y dentro de este *D. caudatum* es la especie que antes diverge, reforzando el carácter relictual de la especie. El origen del linaje de *D. cadatum* se estima que ocurrió durante el Mioceno hace entre 15 y 20 millones de años (Wei et al., 2015). Según la directiva hábitat, a nivel Europeo, *D. caudatum* se encuentra clasificada como en peligro de extinción (Directiva Hábitat, 1992). A nivel estatal, según la Lista Roja de la Flora Vascular Española (Bañares et al., 2010), está clasificada como vulnerable y según el Catálogo Nacional de Especies Amenazadas (Anónimo, 2011) como en peligro de extinción. No existen trabajos previos relacionados con la diversidad genética o filogeografía de esta especie.

4- Pteris incompleta

Pteris incompleta es un helecho que pertenece a la familia Pteridaceae. Se caracteriza por tener un rizoma rastrero, robusto y muy corto, y las frondes pueden alcanzar los 150 cm de altura. Los soros están dispuestos linealmente en los márgenes de la pínula desde la parte basal y no sobrepasando la mitad de ésta. Es una especie diploide 2n=58 (Delgado & Plaza, 2006). Presenta una distribución geográfica y requerimientos ecológicos muy similares a *D. caudatum*. Es un endemismo Europeo-Macaronésico considerado relicto del Terciario que se desarrolla en barrancos profundos, muy umbríos y con una alta humedad atmosférica permanente. Requieren una temperatura elevada y regular, sobre un suelo muy húmedo. Se suele encontrar en arroyos encajados con vegetación de tipo lauroide y abundantes nieblas (Delgado & Plaza, 2006). Actualmente se encuentra como relicto en los archipiélagos macaronésicos Azores, Islas Canarias (Gran Canaria, La Gomera, Tenerife y La Palma) y Madeira y en el sur de la Península Ibérica, concretamente, las Sierras de Algeciras (Cádiz). A nivel estatal, según la Lista Roja de la Flora Vascular Española (Bañares et al., 2010) está clasificada como vulnerable y según el Catálogo Nacional de Especies Amenazadas (Anónimo, 2011) como en peligro de extinción. No existen trabajos previos relacionados con la diversidad genética o filogeografia de esta especie.

El objetivo principal de esta tesis es:

-Analizar la diversidad genética y el patrón filogeográfico de *Vandenboschia speciosa*, *Diplazium caudatum*, *Pteris incompleta* y *Culcita macrocarpa* considerando el área completa de su rango de distribución.

Referencias

Amigo JA, & Norman G. 1995. Identification of site-types important for rare ferns in an area of deciduous woodland in northwest Spain. *Vegetatio* 116: 133-146.

Anonymous. 1979. Convention of the Conservation of European Wildlife and Natural Habitats. *European Treaty Series* 104.

Anonymous. 1992. Council Directive 92/43/EEC of 21 May 1992 on the conservation of habitats and wild fauna and flora. *Official Journal of the European Communities L* 206: 7–50.

Axelrod DI. 1975. Evolution and biogeography of Madrean-Tethyan sclerophyll vegetation. *Annals of the Missouri Botanical Garden* 62: 280–334.

Avise JC. 1994. *Molecular Markers, Natural History and Evolution*. Chapman & Hall. New York, USA.

Bañares Á, Blanca G, Güemes J, Moreno JC, & Ortiz, S. (eds). 2004. *Atlas y Libro Rojo de la Flora Vascular Amenazada de España*. Dirección General de Conservación de la Naturaleza, Madrid.

Barrón E. 2003. Evolución de las floras terciarias en la Península Ibérica. *Monografías del Jardín Botánico de Córdoba,* 11: 63–74.

Barrón E, Peyrot D. 2006. La vegetación forestal en el Terciario. In: Carrión J, Fernández S, Fuentes N, eds. *Paleoambientes y cambio climático*. Murcia, Spain: Fundación Séneca/Agencia de Ciencia y Tecnología de la Región de Murcia, 56–77.

Bystriakova N, Ansell SW, Russell SJ, Grundmann M, Vogel JC, Schneider H. 2014. Present, past and future of the European rock fern *Asplenium fontanum*: combining distribution modelling and population genetics to study the effect of climate change on geographic range and genetic diversity. *Annals of Botany* 113: 453–465.

Chambers GK, & MacAvoy ES. 2000. Microsatellites: consensus and controversy. *Comparative Biochemistry Physiology* 126: 455–476.

Cheng HH, & Crittenden LB. 1994. Microsatellite markers for genetic mapping in the chicken. *Poultry Science* 73: 539–546.

Delgado AJ, & Plaza L. 2006. *Helechos amenazados de Andalucía*.Consejería de Medio Ambiente. Junta de Andalucía. Sevilla.

Domínguez–Domínguez O, & Vázquez E. 2009. Filogeografía: aplicaciones en taxonomía y conservación. *Animal Biodiversity and Conservation* 32: 59–70.

Ellegren H. 2004. Microsatellites: simple sequences with complex evolution. *Nature reviews genetics* 5: 435.

Fernández-Palacios JM, de Nascimento L, Otto R, Delgado JD, García-del-Rey E, Arévalo JR, Whittaker RJ. 2011. A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography* 38: 226–246.

Fernandez-Silva Iria, et al. 2013. Microsatellites for next-generation ecologists: a post-sequencing bioinformatics pipeline. *PLoS one* 8: e55990.

Ferreras C. & Arozena ME. 1987. *Guía Física de España. Los bosques*. Alianza Editorial, Madrid, pp. 394.

Hewitt GM. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* 68: 87–112.

de Groot GA, Verduyn B, Wubs EJ, Erkens RH, During HJ. 2012. Inter-and intraspecific variation in fern mating systems after long-distance colonization: the importance of selfing. *BMC Plant Biology* 12: 3.

Hunt HV, Ansell SW, Russell SJ, Schneider H, Vogel JC. 2009. Genetic diversity and phylogeography in two diploid ferns, *Asplenium fontanum* subsp. *fontanum* and *A. petrarchae* subsp. *bivalens*, in the western Mediterranean. *Molecular Ecology* 18: 4940–4954

Jarne P, & Lagoda PJ. 1996. Microsatellites, from molecules to populations and back. *Trends in ecology & evolution* 11: 424-429.

Lanteri AA, & Confalonieri VA. 2003. Filogeografía: objetivos, métodos y ejemplos. *Una perspectiva Latinoamericana de la Biogeografía (JJ Morrone and J Llorente, eds.). CONABIO, Mexico*, 185-194.

Li YC, Korol AB, Fahima T, Beiles A, Nevo E. 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular ecology* 11: 2453-2465.

Maccagni A, Parisod C, Grant JR. 2017. Phylogeography of the moonwort fern *Botrychium lunaria* (Ophioglossaceae) based on chloroplast DNA in the Central-European Mountain System. *Alpine Botany* 127: 185–196.

Mai DH. 1989. Development and regional differentiation of the European vegetation during the Tertiary. *Plant Systematics and Evolution* 162: 79–91.

Mai DH. 1991. Palaeofloristics changes in Europe and the confirmation of the Arctotertiary–Palaeotropical geofloral concept. *Review of Palaoeobotany and Palynology* 68: 29–36.

Manton I, Lovis JD, Vida G, Gibby M. 1986. Cytology of the fern flora of Madeira. *Bulletin British Museum (Natural History). Botany* 15: 123-161.

Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M. 1997. Genome mapping, molecular markers and marker-assited selection in crop plants. *Molecular Breeding* 3: 87-103.

Moya P, Delgado A, Plaza L. 2010. Propagación y aclimatación de cuatro especies de Pteridófitos amenazados en la provincia de Cádiz 13.

Postigo Mijarra JM, Barrón E, Gómez Manzaneque F, Morla C. 2009. Floristic changes in the Iberian Peninsula and Balearic Islands. *Journal of Biogeography* 36: 2025–2043.

Provan J, & Bennett KD. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution* 23: 564–571.

Quintanilla LG, Pangua E, Amigo J, Pajarón S. 2005. Comparative study of the sympatric ferns Culcita macrocarpa and Woodwardia radicans: sexual phenotype. *Flora-Morphology, Distribution, Functional Ecology of Plant*s 200: 187–194.

Quintanilla LG, Pajarón S, Pangua E, Amigo J. 2007. Allozyme variation in the sympatric ferns Culcita macrocarpa and Woodwardia radicans at the northern extreme of their ranges. *Plant Systematics and Evolution* 263: 135–144.

Rodríguez-Sánchez F, Arroyo J. 2008. Reconstructing the demise of Tethyan plants: climate-driven range dynamics of *Laurus* since the Pliocene. *Global Ecology and Biogeography* 17: 685–95.

Rumsey FJ, Vogel JC, Russell SJ, Barrett JA, Gibby M. 1999. Population structure and conservation biology of the endangered fern *Trichomanes speciosum* Willd. (Hymenophyllaceae) at its northern distributional limit. *Biological Journal of the Linnean Society* 66: 333–344.

Rumsey FJ, Barrett JA, Gibby M, Russell SJ, Vogel JC. 2005. Reproductive strategies and population structure in the endangered pteridophyte *Trichomanes speciosum* (Hymenophyllaceae: Pteridophyta). *Fern Gazette* 17: 205.

Rumsey FJ, Jermy AC, Sheffield E. 1998. The independent gametophytic stage of *Trichomanes speciosum* Willd. (Hymenophyllaceae), the Killarney Fern and its distribution in the British Isles. *Watsonia* 22: 1–20.

Stewart JR, Lister AM. 2001. Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology & Evolution* 16: 608–613.

Suter M, JJ Schneller, & JC Vogel. 2000. Investigations into the genetic variation, population structure, and breeding systems of the fern *Asplenium trichomanes* subsp. *quadrivalens*. *International Journal of Plant Sciences* 161: 233– 244.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cossons JF. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* 7: 453– 464.

Trewick SA, Morgan-Richards M, Russell SJ, Henderson S, Rumsey FJ, Pintér I, Barret JA, Gibby M, Vogel JC. 2002. Polyploidy, phylogeography and Pleistocene refugia of the rockfern *Asplenium ceterach*: evidence from chloroplast DNA. *Molecular Ecology* 11: 2003–2012.

Wei R, Xiang Q, Schneider H, Sundue MA, Kessler M, Kamau PW, Hidayat A, Zhang X, 2015. Eurasian origin, boreotropical migration and transoceanic dispersal in the pantropical fern genus *Diplazium* (Athyriaceae). *Journal of Biogeography* 42: 1809–1819. https://doi.org/10.1111/jbi.12551

Widmer A, & Lexer C. 2001. Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends in Ecology & Evolution* 16: 267–269.

Chapter 1. Genetic diversity and population history of the Killarney fern, *Vandenboschia speciosa* (Hymenophyllaceae), at its southern distribution limit in continental Europe.

Samira Ben-Menni Schuler, María del Carmen García-López, Inmaculada López-Flores, Marta Nieto-Lugilde y Víctor N. Suárez-Santiago

Botanical Journal of the Linnean Society. 183 (2017), 94-105.

Abstract

The improvement of the conservation and management strategies of endangered species requires an understanding of the history and current state of their populations, and thus population genetic studies become essential. This is especially critical for the relict populations at the low-latitude limit of the species range, which have persisted at suitable growing sites across Tertiary and Quaternary climatic changes and are expected to have retained a high proportion of the genetic diversity of the species. Here we investigate the genetic diversity and structure of *Vandenboschia speciosa*, a vulnerable endemic European-Macaronesian species with disjunct distribution, at its southernmost distribution limit in continental Europe (S Iberian Peninsula: Andalusia, Spain). We analysed six sporophyte populations from the 8 known populations using 10 microsatellite loci. The results indicated moderate genetic diversity at the regional scale but low at the population level, and a significant genetic differentiation between populations. Coalescent analyses revealed a population history in Andalusia characterized by a migration-drift equilibrium with historical dispersals as the main factors influencing population structure, but now predominately influenced by genetic drift, according to the genetic differentiation detected. Both spatial autocorrelation and genetic-structure analyses (multivariate and STRUCTURE analyses) suggest that historic gene flow was mainly over short distances, since significant genetic relationships were detected between pairs of nearby populations (< 7 km). Just one population, the most diverse, evidences long-distance dispersal. This study provides a baseline for designing appropriate conservation strategies of *V. speciosa* in Andalusia, where the regional government has recently approved a recovery and conservation plan.

Additional Keywords: Genetic diversity – microsatellites – migration-drift equilibrium – plant conservation – rear-edge populations.

1. Introduction

The presence of subtropical and warm-temperate species in Western Europe is explained by the existence of regions that have acted as climatic refugia, providing suitable conditions for species during both the Tertiary tropical-subtropical climate deterioration and the Pleistocene glacial period (Jermy, 1984; Svenning, 2003; Calleja, Benito-Garzón & Sáinz-Ollero, 2009). Many Tertiary relicts persist in southern Europe (cf. Calleja *et al.*, 2009), especially in the southern peninsulas where most European plant species were constrained during the ice ages (Bennett, Tzedakis & Willis, 1991; Hewitt, 2004).

Stable relict populations at the low-latitude margins of species' distribution ranges, called rear-edge populations, usually harbour the majority of the genetic diversity of the species (Petit *et al.*, 2003; Hewitt, 2004). Frequently these populations are small, restricted to microhabitats within unsuitable landscapes, and have undergone long isolation. Consequently they often show low within-population genetic diversity, but high levels of inter-population genetic differentiation, leading to high levels of regional genetic diversity; moreover, they preserve a high genetic distinctiveness with regard other populations of the species (cf. Hampe & Petit, 2005). Conservation of the rear-edge populations is a priority for conservation of the genetic diversity, phylogenetic history and evolutionary potential of the species (Hampe & Petit, 2005).

Vandenboschia speciosa (Willd.) G. Kunkel (syn. *Trichomanes speciosum* Willd., the Killarney fern; Hymenophyllaceae) is the only European representative of a predominantly tropical genus, being a Macaronesian-European endemic fern (Rumsey, Jermy & Sheffield, 1998a). It has a peculiar life cycle among ferns, in which both phases (sporophyte and gametophyte) are perennial and capable of vegetative propagation. Therefore both mixed and independent sporophyte and gametophyte populations are usual throughout the distribution of the species (Rumsey *et al.*, 1998b). The populations with sporophytes are restricted to the European Atlantic coast and the Macaronesian Islands, in places considered refugia, while independent gametophyte populations occur also in continental Northern Europe and Central Europe (Rumsey *et al.*, 2005 and references therein). The rarity of *V. speciosa* and its disjunct distribution has led it to be considered one of the most vulnerable plant species in continental Europe, being protected under both the Bern Convention and the Habitats Directive (Anonymous, 1979, 1992).

In continental Europe the southernmost populations are those in the Aljibe Mountains at the southern tip of the Iberian Peninsula (Fig. 1), on the northern side of the Strait of Gibraltar (provinces of Cádiz and Málaga, Andalusia, Spain), an area considered

a refugium for the species (Rumsey *et al.*, 2005) and others elements considered Tertiary relicts (Rivas-Goday, 1968; Ojeda, Marañón & Arroyo, 2000; Arroyo *et al.*, 2001; Mejías, Arroyo & Marañón, 2007; Moreno-Saiz & Lobo, 2008). The area has been protected since 1989 as *Los Alcornocales* Natural Park.

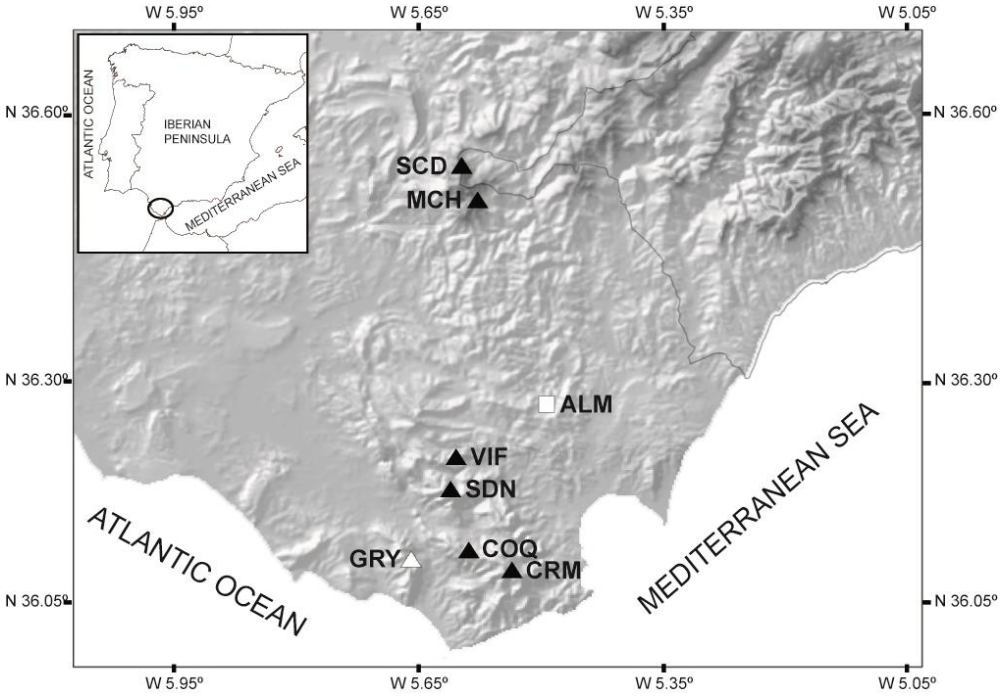


Figure 1. Distribution of the 8 known populations of *V. speciosa* in Andalusia. Triangles represent sporophyte (or mixed) populations, while the square is the only independent-gametophyte population. Black triangles are the populations sampled for this study. Population codes as in Table 1.

Currently 8 populations are known in the area (Fig. 1), with a total of 301 sporophyte individuals (Anonymous, 2015). In 2004 the Andalusia government started a conservation plan (Proyecto de Conservación de Pteridófitos en Andalucía) aimed at monitoring and evaluating the different localities inventoried of several endangered species, including *V. speciosa*. The census established in 2007 was 7 populations, one composed only of the gametophyte phase of the species and a total of 181 sporophyte individuals. During sampling conducted for this study in 2008 the same populations were registered and a total of 161 sporophytes were observed, noting the disappearance of 20 individuals (Suárez-Santiago, Blanca-López & Romero-García, 2010). Only recently, with the continuation of exploration works a new small population has been located as well as a new nucleus with many individuals in one of the populations previously identified.

The *V. speciosa* populations in the Aljibe Mountains meet the criteria for relict rear-edge populations: small population size, a low number of populations persisting in

microhabitats, geographical isolation among populations, and high isolation with regard to other Macaronesian-European populations. Therefore, they are vulnerable to stochastic demographic or climatic events as well as to inbreeding depression, what need to be taken into account by conservation strategies. However, given that these relict populations have persisted in isolation over the Tertiary and Quaternary climatic oscillation, it is possible that demographic stochasticity had played a comparatively minor role in the extinction process (Hampe & Petit, 2005), another aspect to be taken into consideration by conservation strategies. Therefore, improved conservation and management strategies of *V. speciosa* populations in Andalusia require an understanding of the history and current state of these populations, for which population genetic studies are essential.

Only one previous study on the genetic diversity of the Andalusian populations is available, covering three populations and using allozymes as molecular markers (four enzyme systems) coded as phenotypes (Rumsey *et al.*, 2005). In that study the authors detected genetic diversity within populations (without quantifying it) but, based on the detected private alleles and unique multilocus phenotypes, they suggested low levels of gene flow even between intra-population nuclei; moreover, they proposed that inter-gametophyte mating occurs in the study populations on the basis of the observed allele patterns. Therefore, this early study seems to fit the expected model for rear-edge populations.

In this study, using species-specific microsatellite markers, we conducted an exhaustive population genetic study of the Andalusian populations of *V. speciosa*. The aims are (1) to estimate the current levels of genetic diversity and population structure in this southern refugium, (2) to evaluate the correlation of the spatial structure of the genetic diversity with regard to the landscape, and (3) to infer the historical demography of populations in order to know the extent of stochastic demographic events on their persistence. These data provide a baseline for designing appropriate conservation strategies of *V. speciosa* in Andalusia, where an action plan specifying the measures to be taken in the recovery and conservation plan has been recently approved by the Andalusia government (Anonymous, 2015).

2. Materials and Methods

2.1. Plant material

A total of 6 sporophyte populations were sampled. These comprised all the sporophyte populations known in Andalusia at the time of the sampling and until the recent discovery of a new small population and a new intra-population nucleus of an existing population. Table 1 includes data from each population sampled and Fig. 1 shows the location of the 6 populations and the topographical features of the region studied. For each population, frond samples were collected from all adult individuals (except those in inaccessible places; i.e. a total of 99 of the 161 existing individuals).

Table 1. Sampling details of *V. speciosa* populations used in the present study.

Population	Location	Voucher	Geographic coordinates	Altitude (m a.s.l.)	Sample size*
COQ	Canuto de Ojén Quesada (Cádiz: Los Barrios)	GDA 61589	N36.127º/W5.585º	311	56/84
CRM	Cabecera del río de la Miel (Cádiz: Algeciras)	GDA 62522	N36.105°/W5.528°	561	7/13
MCH	Moracha (Cádiz: Alcalá de los Gazules)	GDA 62523	N36.497°/W5.584°	605	20/34
SCD	Garganta de la Sauceda (Málaga: Cortes de la Frontera)	GDA 62524	N36.535°/W5.605°	704	3/3
SDN	Sierra del Niño (Cádiz: Los Barrios)	GDA 62525	N36.186°/W5.610°	560	4/4
VIF	Valdeinfierno (Cádiz: Los Barrios)	GDA 62526	N36.224°/W5.604°	244	9/21

GDA, herbarium of the University of Granada (Spain). *Sampled individuals/total number of individuals observed when sampling was made in 2008.

2.2. DNA extraction and genotyping

Total genomic DNA of the 99 sampled individuals was extracted from fresh fronds using the kit NucleoSpin Plant II (Machery-Nagel), following the manufacturer's instructions. All the individuals were genotyped for 10 recently developed microsatellite loci as described by García-López *et al.* (2015). Genotyping was performed on an ABI PRISM® 3100-Avant Genetic Analyzer (Applied Biosystems) at the Centro de Instrumentación Científica (CIC) of the University of Granada. Alleles were scored using GENEMARKER v1.85 (SoftGenetics). Although none of the 99 individuals sampled showed more than two alleles per locus, all individuals were treated as tetraploid based on previous studies (Manton, 1950, 1986; Mehra & Singh, 1957; Obermayer *et al.*, 2002) and the data we are compiling from genotyping gametophytes in populations studied here (unpubl. data), where many of them are diploid (heterozygous for two different alleles).

The MAC-PR method (Esselink *et al*., 2004) was used to obtain the allele-dosage for each individual in the analysed loci (see Supporting Information, Fig. S1). Proper assignments of the alleles into tetraploid configurations were clear in most of the cases; in the few cases where assignment was ambiguous, the configuration was completed with alleles scored as 'missing data' (6 individuals: 1 for locus VS-CA1-55, 2 for VS-HKPD7, and 3 for VS-HM60O).

2.3. Statistical analysis

Genetic diversity indices [number and average number of alleles, observed and unbiased expected heterozygosity (Nei, 1987), and inbreeding coefficient (Weir & Cockerham, 1984)] were calculated with GenoDive version 2.0b24 using, when necessary, the maximum likelihood method to correct for the unknown dosage of the alleles in the individuals with missing data (Meirmans & Van Tienderen, 2004), and considering only the unique multilocus genotypes (genets). Identical multilocus genotypes (possible clones) were previously identified using the Assign Clones function in GenoDive.

The existence of genetic differentiation between populations was assessed by the fixation index (G_{ST}; Nei, 1987), calculated with GenoDive and using the maximum likelihood method to correct for the unknown dosage of the alleles; the significance of G_{ST} was tested by a permutation test with 10,000 permutations. Distribution of genetic variability within and among populations was evaluated using an analysis of molecular variance (AMOVA; Excoffier *et al*., 1992) and tested with a permutation test (10,000 permutations) with GenoDive. The population genetic structure was analysed using two different approaches. On one hand, the Bayesian algorithm implemented in STRUCTURE v2.3.3 (Falush, Stephens & Pritchard, 2003) was used to evaluate the number of genetic clusters (*K*) in our data. We tested the number of clusters ranging from 1 through 8, with 10 replicates per *K*, using the admixture model and independent allele frequencies. We used independent frequencies because of the long isolation between populations we expected, on the base of Rumsey *et al.* 2005. The burn-in period was set to 50,000 and Markov Chain Monte Carlo iterations set to 10^6. The number of clusters was assayed using the Delta *K*-method (ΔK) proposed by Evanno, Regnaut & Goudet (2005), using the online tool Structure Harvester v. 0.6.94 (Earl & vonHoldt, 2012). To align and visualize the STRUCTURE output across the 10 replicates, we used the online program CLUMPAK (Kopelman *et al.*, 2015). On the other hand, the genetic structure was described using two multivariate-statistics-based clustering methods, a principal

coordinate analysis (PCoA; Gower, 1966) and K-means clustering (Hartigan & Wong, 1979). As the multivariate analysis lacks underlying population genetics assumptions, then they are independent of the ploidy level, and they are therefore suitable to polyploid or mixed-ploidy analyses. The PCoA was performed with the R-package Polysat v. 1.4-1 (Clark & Jasienuk, 2011), using the function *meandistance.matrix* to estimate the Bruvo distance (Bruvo *et al.*, 2004) between pairs of individuals and the *cmdscale* function to carry out the PCoA on the distance matrix. The K-means clustering was performed with GenoDive using the simulated annealing method with 50,000 steps and 500 random starts, and using the maximum likelihood method in order to correct for the unknown dosage of alleles; for this analysis the Bruvo distance matrix previously calculated with Polysat was used. The clustering was performed for a range of *K* values from 2 to 15, considering the optimal clustering to be the one with the highest value for the pseudo-F-statistic (Calinski-Harabasz, 1974).

The 2MOD v. 0.2 program was used to investigate the population history of *V. speciosa* in Andalusia on the basis of the coalescent theory (Ciofi *et al.*, 1999). The method compares the relative likelihoods for a model of migration-drift equilibrium (gene flow model) vs. ancestral population fragmentation into independent units diverging by drift (drift model). The F-value (probability that two alleles share a common ancestor within a population) for each population and the number of migrants per generation [M= (1-F)/(4F); Ciofi *et al.*, 1999] were also estimated. The program was run with a Markov chain Monte Carlo simulation of 1,000,000 iterations; the 10% of the output was discarded as the burn-in period.

Isolation by distance was tested in the 6 populations using regression of pairwise G'_{ST} distances (Nei, 1987; determined with GenoDive using the polyploidy dosage correction, and transformed as $G'_{ST}/(1-G'_{ST})$ and logarithms of geographic distances between populations, using a Mantel test in GenoDive. Moreover, GenoDive was also used to perform a spatial autocorrelation analysis (using the transformed G'_{ST} distances) to test the genetic similarity between populations considering three geographic-distance classes and then to understand how far dispersions occur. The geographic-distance classes were (see Supporting Information, Fig. S2): short (the upper limit being the maximum distance between neighbouring populations; 6,850 m), medium (distance between two populations with one or two population in between, within each two main spatial clusters of populations, and with the upper limit being 15,000 m), long (distance higher than

15,000 m). Significance of both analyses (isolation by distance and spatial autocorrelation) was tested by 10,000 permutations.

3. Results

One of the 10 loci tested was monomorphic (VS-GT0D8-1) and therefore not considered for the analyses. All 99 individuals showed only one or two alleles per locus. In total, 84 multilocus genotypes (MLG) were found among the 99 individuals sampled (Table 2). At the intra-population level, all populations had some MLG shared between two or more individuals, but no MLG was shared between populations.

Table 2. Summary of genetic variation for 9 microsatellite loci and test for migration-drift equilibrium in the 6 Andalusian studied populations of *V. speciosa*.

							Migration-drift test	
Population[a]	N	A	Priv	H_O	H_E	F_{IS}[b]	F	M
COQ	53/56	31(3.44)	10	0.187	0.327	0.496*	0.320	0.53
CRM	5/7	26(2.89)	4	0.285	0.387	0.424*	0.109	2.05
MCH	14/20	17(1.89)	1	0.139	0.164	0.322*	0.720	0.10
SCD	2/3	11(1.22)	1	0.074	0.068	0.333	0.857	0.04
SDN	3/4	18(2.00)	3	0.310	0.375	0.272*	0.502	0.25
VIF	6/9	13(1.44)	2	0.077	0.091	0.336*	0.783	0.07
Total/mean	83/99	49(5.44)	21	0.179	0.239	0.458*	0.548	0.415

N = number of genets/number of ramets; A = total number and mean number (in parenthesis) of alleles per locus; Priv: number of private alleles; H_O = observed heterozygosity; H_E = unbiased expected heterozygosity (Nei, 1987); F_{IS} = inbreeding coefficient (Weir & Cockerham, 1984; * = $P < 0.01$); F = probabilities of allelic co-ancestory; M = number of migrants per generation. [a]Population codes as in Table 1. [b]F_{IS} calculated using H_E values obtained without correction for the unknown dosage of the alleles.

A total of 49 alleles were observed from the 9 loci surveyed (average per locus: 5.44). The total genetic diversity (H_T) for *V. speciosa* in Andalusia (considering only the sporophytes) was 0.545 (SD: 0.042). At the population level (Table 2), the mean number of alleles observed ranged from 1.22 in SCD to 3.44 in COQ (mean per population: 2.15). The lowest value of genetic diversity was found in SCD (H_E = 0.068) whereas the highest value was found in CRM (H_E = 0.387); the mean genetic diversity value per population was 0.239 (Table 2). Both the SCD and VIF populations showed a very high proportion of monomorphic loci (7 and 5, respectively). Population exclusive alleles (21 in total; Table 2) were detected in all populations, highlighting the 10 found in COQ. It is also noteworthy that MCH, the second largest population, showed only one private allele.

Virtually all populations showed statistically significant high positive values of inbreeding coefficients, except SCD population in which statistical significance was not reached.

The genetic differentiation between populations was highly significant (G_{ST} = 0.561 ± 0.05; $P < 0.001$). 47.1% of genetic variation was found among populations (F_{ST} = 0.471; $P < 0.001$) while 24.1% and 28.8% was among individuals within populations and within individuals, respectively (Table 3).

Table 3. Summary of the analysis of molecular variance (AMOVA) of the 6 studied populations of *V. speciosa*.

Source of Variation	Variation (%)	F-value	c.i. 2.5%	c.i. 97.5%	*P*
Within Individual	28.8	0.712	0.626	0.807	-
Among Individual	24.1	0.456	0.402	0.522	0.000
Among Population	47.1	0.471	0.301	0.651	0.000

c.i. = 95% confidence intervals of F-statistics (obtained through bootstrapping over loci).

The PCoA analysis based on the Bruvo distance assigned almost all individuals to three groups (I-III; Fig. 2a). Thus, individuals from MCH and SCD formed the Group-I, those of COQ formed the Group-II, and all individuals from SDN and VIF where included in Group-III. Individuals from one CRM population belonged to the Group-I, another to the Group-II, and the three remaining occupy an intermediate position between the Groups I and II (Fig. 2a). The K-means clustering analysis mostly confirmed the groupings observed in the PCoA (Fig. 2b). The best clustering according to pseudo-F-statistic was K = 3, and the only difference between this analysis and the PCoA analysis was the assignation to the Group I of the three CRM individuals ambiguously located in the PCoA.

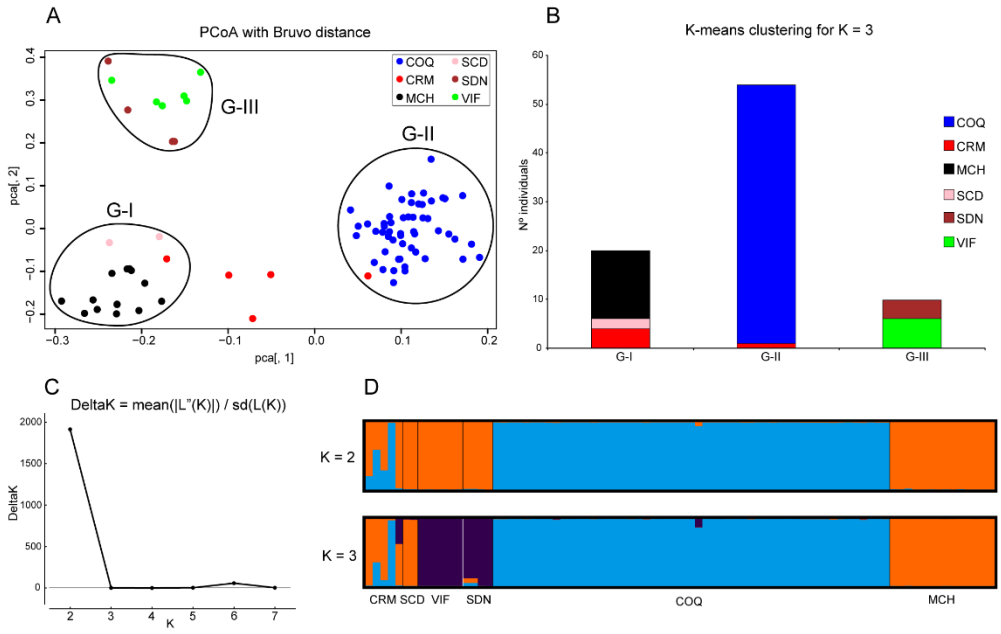


Figure 2. Genetic structure of *V. speciosa* in Andalusia as inferred by multivariate (A, B) and STRUCTURE (C, D) analyses. (A) Result of the principal coordinate analysis (PCoA) using the Bruvo distance between pairs of individuals, showing the three resulting main groups (G-I to G-III). (B) Result of the k-mean clustering analysis for the best clustering according to pseudo-F-statistic, K = 3, showing a histogram of individual assignment to the clusters (G-I to G-III). (C) ΔK values used to select the most probable cluster structuring according to the method proposed by Evanno *et al.* (2005). (D) Histograms of individual assignment to clusters for the two most probable structuring resulting of the STRUCTURE analysis, K = 2 and K = 3. Population codes as in Table 1.

The assignation test implemented by STRUCTURE recognized two genetic clusters (K = 2) as the most probable structuring. One cluster (CL1) was composed by four populations with high mean assignation percentages of their individuals to the cluster (> 99.4%; MCH, SCD, SDN, VIF), and the other cluster (CL2) composed of all COQ individuals with a mean assignation percentage of 99.5%; CRM individuals were assigned to both clusters, two of which were clearly assigned with high probabilities (each to a cluster) and the three remaining with high mixing proportion (Fig. 2d). On considering three genetic clusters (K = 3), as identified in the multivariate analyses, a substructure was detected within CL1 (when K = 2), splitting SDN/VIF from MCH/SCD. The structuring for K = 3 ratified the clustering found in the multivariate analyses, with the same populations forming the three clusters and the CRM individuals divided into CL1 and CL2 (Groups I and II in the multivariate analyses, respectively).

The analysis implemented in the 2MOD program revealed that the Andalusian populations of *V. speciosa* significantly fit a gene flow model (*P* = 1). The F-values

ranged from 0.109 to 0.857 (mean: 0.548), and the number of migrants per generation varied from 0.04 to 2.05 (mean: 0.506) (Table 2).

The Mantel test showed a lack of isolation by distance across the populations (r = 0.244, *P* = 0.119). The spatial autocorrelation analysis showed a significant genetic similarity between neighbouring populations (short distance class, < 7 km; r = 0.343, *P* = 0.045), while the genetic similarity was not significant among populations at medium (7-15 km; r = -0.033, *P* = 0.514) and long distances (15-50 km; r = -0.277, *P* = 0.198).

4. Discussion

This study presents a detailed analysis of the genetic diversity and structure of the southernmost populations of *V. speciosa* on the European continent. It includes 6 of the 7 known sporophyte populations so far known in the region and all adult individuals in each population (except those inaccessible and the new recently discovered in COQ). All 6 populations analysed inhabit the Aljibe Mountains (Cádiz and Málaga provinces), a locally wet habitat in a regionally Mediterranean climate characterized by dry summers. These populations are considered climate relicts that have persisted since the last glacial period or even longer (Rumsey *et al.*, 2005; Delgado-Vázquez & Plaza-Arregui, 2010). Our results reveal the demographic history of *V. speciosa* in the region, in which the gene flow at short distances and population isolation have played a key role.

4.1. Genetic variation, ploidy level and mating system of the andalusian *V. speciosa* populations

Vandenboschia speciosa has largely been considered a tetraploid species on the basis of the Manton (1950), Mehra & Singh (1957), Manton *et al.* (1986), and later Obermayer *et al.* (2002) studies. However, recently Ní Dhúill *et al.* (2015) found populations in Ireland that proved to be a mix of individuals having different ploidy levels (diploid, triploid, and tetraploid sporophytes as well as haploid and diploid gametophytes). Our results from the microsatellite genotyping suggest that the Andalusian sporophytes are diploid, since no individuals show more than two alleles per locus. However, the genotyping of Andalusian gametophytes in an ongoing study (unpubl. data) has revealed that most gametophytes are diploid (most of these with two alleles for at least some loci), and therefore the expected ploidy for sporophytes would be tetraploid. Detecting only two alleles per locus in tetraploid organisms could be explained by disomic inheritance, where a chromosome only pairs with its homolog in an

allotetraploid species. However, although *V. speciosa* is believed to be an allotetraploid species (Ebihara *et al*., 2007), the allelic pattern observed in our study does not conform to a disomic inheritance. Disomy leads to fixed heterozygosity and therefore negative values of F_{IS} (cf. Meirmans and van Tienderen, 2013), while our results show high values of F_{IS} due to the excess of homozygotes even for high population sizes (Table 2). Another possible explanation for the observed allelic pattern is a high incidence of intra-gametophytic mating in the populations. Intra-gametophytic mating in diploids leads homozygous sporophytes at all loci (Lloyd, 1974), but in tetraploids it can produce both homozygous and two-allele heterozygous sporophytes.

Our genetic diversity analyses showed a moderate level of diversity at regional scale (49 alleles in total, mean per locus: 5.44; H_T: 0.545), but the genetic variation was low within each population (alleles per locus: 1.22-3.44, alleles per locus population mean: 2.15; H_E: 0.068-0.387, H_E population mean: 0.239), being CRM, COQ and SDN the most variable, and SCD and VIF the less variable populations. The low genetic diversity detected in VIF could have been underestimated since we could sample only 9 from the 21 individuals observed. Our results agree with the intra-population genetic variation reported by Rumsey *et al.* (2005) in three Andalusian populations (two of these included in our study, COQ and CRM) using allozymes coded as phenotypes. However, the findings here contrast with the lack of intra-population genetic diversity detected in the Scottish (no intra-population variation) and Central European (four of 35 study sites showed more than one multilocus phenotype) populations (Rumsey *et al.*, 1996, 1998b, 1999). Unfortunately, our results on the quantification of the genetic diversity cannot be compared with those of Rumsey *et al.* (2005) because these authors did not use population statistics for quantification. In addition to the existence of intra-population genetic variation, Rumsey *et al.* (2005) found different private alleles for each of the three Andalusian populations they analysed. According to our results, all six populations analysed show private alleles (especially COQ with 10 private alleles), even the low-individual populations. Our study, which includes a wider sampling, supports the conclusion drawn by Rumsey *et al.* (2005) that the southern Iberian Peninsula represents a climatic refugium for *V. speciosa*. In addition, our genetic diversity results support the Andalusian populations of the species meet the criteria for relict rear-edge populations; with low intra-population genetic diversity in contrast to the increased levels of regional genetic diversity. The genetic diversity found for a population of *V. speciosa* in the Basque Country (Spain) by García-López *et al.* (2015; study describing the microsatellite

loci used here) was even higher (H_E: 0.523) than that found for the Andalusian populations in the present study, also suggesting the refugium status of this area (all Cantabrian range) in which *V. speciosa* abounds (Boudrie, 1998; Romero *et al.*, 2005; Sanz-Azkue, Díez-López & Olariaga-Ibarguren, 2013; Sanz-Azkue, Olariaga & Felipe 2014), and which has also been demonstrated for other plant species (e.g., Benito-Garzón, Sánchez de Dios, Sáinz-Ollero, 2007; Calleja *et al.*, 2009; Beatty & Provan, 2012).

The mating system of a species is usually estimated from the genetic structure of its populations, where outcrossing species show genetic variation mainly within populations and selfing species does among populations (cf. Bucharová & Münzbergová, 2012). Regarding *V. speciosa*, intra-population genetic diversity has been considered as evidence of sexual inter-gametophytic mating (Rumsey *et al.*, 2005). Therefore, the genetic diversity detected in the Andalusian populations (although it is low) suggests that inter-gametophytic mating occurs, unlike the scenario suggested for the Scottish and most of the Central European populations (Rumsey *et al.*, 1998b, 1999, 2005). However, our results of AMOVA suggest *V. speciosa* in Andalusia conforms to the expectations for a selfing species, with most genetic variation partitioned among populations. Additionally, the populations show a lack of heterozygotes resulting in high levels of inbreeding. Even in COQ population, which is one of the most variable populations and having a higher number of multilocus genotypes, the inbreeding coefficient was the highest. This pattern suggests that intra-population variation is largely achieved by the entry of new genotypes, which afterwards reproduce mainly by intra-gametophytic mating. A similar pattern was previously observed on the northern colonization front of several rock-dwelling ferns with mixed mating system, as for example *Adiantum capillus-veneris* L. (Pryor *et al*., 2001) and *Asplenium scolopendrium* L. (De Groot *et al*., 2012a). Therefore, although we have detected intra-population genetic diversity, our results seem to support high occurrence of intra-gametophytic mating in the studied populations; which agree with the observed allelic pattern, as we noted before.

4.2. Effect of the gene flow and genetic drive in the genetic structure of the andalusian *V. speciosa* populations

Our results reflect a significant genetic differentiation between the Andalusian populations of *V. speciosa*. Inter-population structure of the genetic diversity has been previously reported in *V. speciosa* using allozyme analysis (Rumsey *et al.* 1998b, 1999), including the three Andalusian populations analysed by Rumsey *et al.* (2005).

Results from the 2MOD analysis strongly support migration-drift equilibrium in Andalusian *V. speciosa* populations, suggesting that gene flow had a key influence on population structure (as expected for a highly dispersible fern). However, the high F-values registered for all populations suggest that they are currently predominately influenced by the genetic drift, according to the genetic differentiation detected, with very low migration values (Table 2). A similar case has been detected in the endemic *Adiantum reniforme* L. var. *sinense* Y.X.Lin in the Three Gorge Reservoir Area (China; Kang *et al.*, 2008), for which the population history is characterized by a balance of gene flow and genetic drift. The many private alleles found in the Andalusian populations of *V. speciosa* support the action of the genetic drift in them (Table 2), because gene flow would obscure these private variants. Allozyme data found by Rumsey *et al.* (2005) in the Andalusian populations also support the genetic drift effect.

Given that *V. speciosa* is a highly dispersible fern, the low migratory rates detected among the Andalusian populations could be related to its colonization dynamics, including the effect that the reproductive system has on this dynamics. Gametophyte of *V. speciosa* is perennial and it reproduces vegetatively in a very efficient way. Moreover, gametophyte can be dispersed by the production of gemmae. Rumsey *et al.* (1998b, 1999), based on these characteristics of the *V. speciosa* gametophyte, proposed the block of niches as a cause favouring the absence of gene flow between the highly differentiated Scottish and also Central European gametophytic populations. According these authors, an earlier arrival of propagules may be able fully to occupy and keep filled a site by vegetative reproduction, excluding later arrivals, in sparsely distributed and locally small suitable habitats. In Andalusia the intra-gametophytic mating suggested by our results would enhance the colonization abilities of the specie, favouring the population establishment from single or few spores (Schneller & Holderegger, 1996; Suter *et al.*, 2000; Lott *et al.*, 2003; Bucharová & Münzbergová, 2012; De Groot *et al.*, 2012a, b), but for the same reason it can also enhance the block of niches proposed by Rumsey *et al.* (1998b, 1999).

Both spatial autocorrelation and genetic-structure analyses (multivariate and STRUCTURE analyses) suggest that historic gene flow was mainly over short distances, since significant genetic relationships were detected between pairs of nearby populations forming patches in short distances (< 7 km). Within each population pair a population with many individuals and another with few individuals were involved (MCH-SCD/VIF-SDN/COQ-CRM), suggesting that the small populations were formed by dispersal from

the larger ones. In the case of CRM the analyses show that it was influenced by COQ, but also by MCH (Fig. 2A, B, D). The results suggest that CRM could have received gene flow from both populations, a situation that would account for CRM being the population with highest H_E and M (Table 2). Gene flow between MCH and CRM implies that long-distance events have happened (44 km).

Our results show the population history of *V. speciosa* in its southernmost distribution area of continental Europe has been modelled by a balance between dispersal frequency at one side (higher at short distance, lower at medium and high distance) and genetic drift and rare establishment opportunities on the other side, with the colonization dynamic of the species affecting such balance. This is in line with patterns found more often in fern populations (Schneller & Holderegger, 1996; Suter *et al*., 2000; Landergott *et al*., 2001; Lott *et al*., 2003; Bucharová & Münzbergová, 2012; De Groot *et al*., 2012a).

4.3. Implications for conservation

In Andalusia, *V. speciosa* is included in the Red List of the Andalusian Vascular Flora as critically endangered (Cabezudo *et al.*, 2005), and it is protected under the Andalusian law of wild flora and fauna (8/2003), where the species is included as vulnerable in the Andalusian catalogue of endangered species. All populations are within the limits of the *Los Alcornocales* Natural Park, but major risk factors nevertheless persist, especially uncontrolled silviculture, alteration of watercourses and deviation of water flows for livestock use (Delgado-Vázquez & Plaza-Arregui, 2010). These threats are affecting the habitat of the species and could diminish the number of suitable niches for individual development. Consequently, we can expect a decline in the number of individuals and greater population isolation.

The low levels of gene flow detected are not sufficient to balance the effect of drift on the genetic diversity of the populations and in the latter case to counteract the negative effect on the population viability (Mills & Allendorf, 1996). Therefore, according our results the measures to be taken should aim at reversing the decline in the number of individuals per population, by translocations, as well as at counteracting population isolation, by fomenting their interconnections. In this sense, and considering that dispersal occurs mainly below 7 km, a first priority should be conserve sufficient patches of this type of habitat less than 7 km from each other, to allow stepping stones for exchange, and that at the same time, key habitat patches that are currently inhabited by populations should not be reduced in size. Recently, the Andalusian government has

approved the Ferns Recovery and Conservation Plan (Anonymous, 2012), included in the Andalusian Strategy of Comprehensive Management of Biodiversity, which includes an Action Program for the years 2015-2019 detailing the measures to be taken in this period (Anonymous, 2015). With respect to *V. speciosa*, the priority actions are: the habitat management, the development of a protocol for *ex situ* production of sporophytes, the reinforcement of populations. Specifically, the reinforcement will be in the COQ and SDN populations, experimentally using gametophytes (because up to now *ex situ* sporophyte propagation has not been successful) produced in the Laboratory of Plant Propagation of the Andalusian government.

The results presented in this paper provide a baseline study for optimal planning of the reinforcement/(re)introduction measures, as they provide a priority in the populations to be reinforced, in the populations to be considered as the sources for translocations, and also in populations that should be considered in order to collect material (spores, gametophytes, sporophytes) for germplasm banks harbouring high genetic variability. Thus, the most diverse and numerous populations (COQ, CRM, MCH) should be prioritized for conservation (preventing the loss of individuals) and for collection of material for germplasm banks that can be used in future reinforcements/(re)introductions. Those small populations representing dispersal events, i.e. SCD and SDN, appear to be the most suitable for reinforcement. However, as a means of reducing the admixture of individuals from genetically distinct populations and avoiding exogamy (Hufford & Mazer, 2003), the genetic groups detected should be taken into account. Thus, MCH and VIF should be the sources for translocation into SCD and SDN, respectively.

In this study, we have detected and quantified the genetic variability within and between *V. speciosa* sporophyte populations, and we have considered the implications for the conservation of this species. However, given the peculiarity of the life cycle of this species (with a perennial gametophyte with an active vegetative reproduction, and the ability to tolerate darker and drier habitats than does the sporophyte; Rumsey & Sheffield, 1996), it seems necessary to include gametophytes in the genetic analyses. On the other hand, the exemption of the gametophyte could lead to an underestimation of the diversity and structure of the populations. These studies should be made before using gametophytes in the conservation plans of *V. speciosa*, so as to be able to properly plan the actions.

Acknowledgements

The authors thank Antonio Delgado and Dr. Gabriel Blanca for help in sample collection and assistance in the field. Special thanks are due to Dr. Carlos García-Verdugo and two anonymous reviewers for valuable comments on the manuscript. This study was supported by the Regional Andalusian Government (project P10-RNM-6198). S.B.S. was funded by a predoctoral grant (F.P.U. program) from the Spanish Government.

References

Anonymous. 1979. Convention of the Conservation of European Wildlife and Natural Habitats. *European Treaty Series* 104.

Anonymous. 1992. Council Directive 92/43/EEC of 21 May 1992 on the conservation of habitats and wild fauna and flora. *Official Journal of the European Communities* L 206: 7–50.

Anonymous. 2012. Acuerdo de 13 de marzo de 2012, del Consejo de Gobierno, por el que se aprueban los planes de recuperación y conservación de determinadas especies silvestres y hábitats protegidos. *Boletín Oficial de la Junta de Andalucía* 60: 164.

Anonymous. 2015. Orden de 20 de mayo de 2015, por la que se aprueban las programas de actuación de los Planes de Recuperación y Conservación de especies catalogadas de Andalucía. *Boletín Oficial de la Junta de Andalucía* 100: 50.

Beatty GE, Provan J. 2012. Post-glacial dispersal, rather than in situ glacial survival, best explains the disjunct distribution of the Lusitanian plant species *Daboecia cantabrica* (Ericaceae). *Journal of Biogeography* 40: 335–344.

Benito-Garzón M, Sánchez de Dios R, Sáinz-Ollero H. 2007. Predictive modelling of tree species distributions on the Iberian Peninsula during the Last Glacial Maximum and Mid-Holocene. *Ecography* 30: 120–134.

Bennett KD, Tzedakis PC, Willis KJ. 1991. Quaternary refugia of north European trees. *Journal of Biogeography* 18: 103–115.

Boudrie M. 1998. Les ptéridophytes du Pays Basque et du nord- ouest de l´espagne; écologie, répartition, protection. *Journal de Botanique de la Société Botanique de France* 5: 43–52.

Bruvo R, Michiels NK, D'Souza TG, Schulenburg H. 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology* 13: 2101–2106.

Bucharová A, Münzbergová Z. 2012. Gene flow among populations of two rare co-occurring fern species differing in ploidy level. *PLoS ONE* 7: e45855.

Cabezudo B, Talavera S, Blanca G, Salazar C, Cueto M, Valdés B, Hernández-Bermejo JE, Herrera CM, Rodríguez-Hiraldo C, Navas D. 2005. *Lista Roja de la flora vascular de Andalucía.* Sevilla: Consejería de Medio Ambiente, Junta de Andalucía.

Calleja JA, Benito-Garzón M, Sáinz-Ollero H. 2009. A Quaternary perspective on the conservation prospects of the Tertiary relict tree Prunus lusitanica L. *Journal of Biogeography* 36: 487–498.

Caliński T, Harabasz J. 1974. A dendrite method for cluster analysis. *Communications in Statistics-Theory and Methods* 3: 1–27.

Ciofi C, Beaumont MA, Swingland IR, Bruford MW. 1999. Genetic divergence and units for conservation in the Komodo dragon Varanus komodoensis. *Proceedings of the Royal Society of London B: Biolocigal Sciences* 266: 2269–2274

Clark LV, Jasieniuk M. 2011. POLYSAT: an R package for polyploid microsatellite analysis. *Molecular Ecology Resources* 11: 562–566.

De Groot GA, During HJ, Ansell SW, Schneider H, Bremer P, Wubs ERJ, Maas JW, Korpelainen H, Erkens RHJ. 2012a. Diverse spore rains and limited local exchange shape fern genetic diversity in a recently created habitat colonized by long-distance dispersal. *Annals of Botany* 109: 965–978.

De Groot GA, Verduyn B, Wubs ERJ, Erkens RHJ, During HJ. 2012b. Inter- and intraspecific variation in fern mating systems after long- distance colonization: the importance of selfing. *BMC Plant Biology* 12: doi:10.1186/1471-2229-12-3.

Delgado-Vázquez AJ, Plaza-Arregui L. 2010. *Helechos amenazados de Andalucía: avances en conservación*. Sevilla: Consejería de Medio Ambiente, Junta de Andalucía.

Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources* 4: 359–361.

Ebihara A, Dubuisson JY, Iwatsuki K, Ito M. 2007. Systematics of *Trichomanes* (Hymenophyllaceae: Pteridophyta), progress and future interests. *Fern Gazette*

18: 53-58.

Esselink GD, Nybom H, Vosman B. 2004. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting – peak ratios) method. Theoretical and Applied Genetics 109*:* 402–408.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611–2620.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes - Application to Human Mitochondrial-DNA Restriction Data. *Genetics* 131: 479–491.

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics,* 164: 1567–1587.

Frankham R, Lees K, Montgomery ME, England PR, Lowe EH, Briscoe DA. 1999. Do population size bottlenecks reduce evolutionary potential?. *Animal Conservation* 2: 255–260.

García-López MC, Ben-Menni Schuler S, López-Flores I, Nieto-Lugilde M, Terrón-Camero L, Mazuecos-Aguilera I, Suárez-Santiago VN. 2015. Development of polymorphic microsatellite markers for the Killarney Fern (*Vandenboschia speciosa*, Hymenophyllaceae). *Applications in Plant Sciences* 3: 1500067.

Gower JC. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 53: 325–338.

Hampe A, Petit RJ. 2005. Conserving biodiversity under climate change: the rear edge matters. *Ecology letters* 8: 461–467.

Hartigan JA, Wong MA. 1979. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28: 100–108.

Hewitt GM. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 359: 183–195.

Hufford KM, Mazer SJ. 2003. Plant ecotypes: genetic differentiation in the age of ecological restoration. *Trends in Ecology & Evolution,* 18, 147–155.

Jermy A.C. 1984. Origin and distribution of Pteridophytes in the Mediterranean area. *Webbia* 38: 397–416.

Kang M, Huang H, Jiang M, Lowe AJ. 2008. Understanding population structure and historical demography in a conservation context: population genetics of an endangered fern. *Diversity and Distributions* 14: 799–807.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* 15: 1179–1191.

Landergott U, Holderegger R, Kozlowski G, Schneller JJ. 2001. Historical bottlenecks decrease genetic diversity in natural populations of *Dryopteris cristata*. *Heredity* 87: 344–355.

Lloyd RM. 1974. Reproductive biology and evolution in Pteridophyta. *Annals of the Missouri Botanical Garden* 61: 318–331.

Lott MS, Volin JC, Pemberton RW, Austin DE. 2003. The reproductive biology of the invasive ferns *Lygodium microphyllum* and *L. japonicum* (Schizaeaceae): Implications for invasive potential. *American Journal of Botany* 90: 1144–1152.

Manton I. 1950. *Problems in cytology and evolution in the Pteridophyta*. Cambridge: Cambridge University Press.

Manton I, Lovis JD, Vida G, Gibby M. 1986. Cytology of the fern flora of Madeira. *Bulletin of the British Museum (Natural History), Botany Series* 15: 123–161.

Marín JA. 2002. La vida en el límite, o cómo las plantas de los" canutos" nos enseñan sobre climas pasados y cómo prevenir el futuro. *Almoraima: revista de estudios campogibraltareños* 27: 157–168.

Medail F, Quezel P. 1997. Hot-spots analysis for conservation of plant biodiversity in the Mediterranean Basin. *Annals of the Missouri Botanical Garden* 84: 112–127.

Mehra PN, Singh G. 1957. Cytology of hymenophyllaceae. *Journal of Genetics* 55: 379–393.

Meirmans PG, Van Tienderen PH. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4: 792–794.

Meirmans PG, Van Tienderen PH. 2013. The effects of inheritance in tetraploids on genetic diversity and population divergence. *Heredity* 110: 131–137.

Mejías JA, Arroyo J, Marañón T. 2007. Ecology and biogeography of plant communities associated with the post Plio-Pleistocene relict *Rhododendron ponticum subsp. baeticum* in southern Spain. *Journal of biogeography* 34: 456–472.

Mills LS, Allendorf FW. 1996. The one-migrant-per-generation rule in conservation and management. *Conservation Biology* 10: 1509–1518.

Moreno-Saiz JC, Lobo JM. 2008. Iberian–Balearic fern regions and their explanatory variables. *Plant Ecology* 198: 149–167.

Nei M. 1987. Genetic distance and molecular phylogeny. In: Ryman N., Utter F. (Eds), *Population Cenetics and Fishery Management*. Seattle: University of Washington Press.

Ní Dhúill E, Smyth N, Waldren S, Lynn D. 2015. *Monitoring methods for the threatened Killarney Fern (Trichomanes speciosum Willd.) in Ireland*. Irish Wildlife Manuals, No. 82. Dublin: National Parks and Wildlife Service, Department of the Arts, Heritage and the Gaeltacht.

Obermayer R, Leitch IJ, Hanson L, Bennett MD. 2002. Nuclear DNA C-values in 30 species double the familial representation in pteridophytes. *Annals of Botany* 90: 209–217.

Ojeda F, Marañón T, Arroyo J. 2000. Plant diversity patterns in the Aljibe Mountains (S. Spain): a comprehensive account. *Biodiversity & Conservation* 9: 1323–1343.

Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, Mohanty A, Müller-Starck G, Demesure-Musch B, Palmé A, Martín JP, Rendell S, Vendramin G. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* 300: 1563–1565.

Pryor KV, Young JE, Rumsey FJ, Edwards KJ, Bruford MW, Rogers HJ. 2001. Diversity, genetic structure and evidence of outcrossing in British populations of the rock fern *Adiantum capillus-veneris* using microsatellites. *Molecular Ecology* 10: 1881–1894.

Rivas-Goday S. 1968. Algunas novedades fitosociológicas de España meridional. *Collectanea Botanica* 7: 997–1031.

Romero MI, Amigo J, Rodríguez-Guitián MA, Díaz-Varela R, Ferreiro da Costa J. 2005. Conservación de la Pteridoflora amenazada en el NW ibérico (Galicia): las especies incluidas en la Directiva Hábitat. *Bulletin de la Société d´Histoire Naturelle de Toulouse* 141: 227–231.

Rumsey FJ, Jermy AC, Sheffield E. 1998a. The independent gametophytic stage of *Trichomanes speciosum* Willd.(Hymenophyllaceae), the Killarney Fern and its distribution in the British Isles. *Watsonia* 22: 1–20.

Rumsey FJ, Russell SJ, Ji J, Barrett JA, Gibby M. 1996. Genetic variation in the endangered filmy fern *Trichomanes speciosum* Willd. In: Camus JM, Gibby M, Johns RF, eds. *Pteridology in Perspective.* Richmond: Royal Botanic Gardens, Kew, 161–165.

Rumsey FJ, Sheffield E. 1996. Inter-generational ecological niche separation and the 'independent gametophyte' phenomenon. In: Camus JM, Gibby M, Johns RF, eds. *Pteridology in Perspective.* Richmond: Royal Botanic Gardens, Kew, 563–570.

Rumsey FJ, Vogel JC, Russell SJ, Barrett JA, Gibby M. 1998b. Climate, colonisation and celibacy: population structure in Central European *Trichomanes speciosum* (Pteridophyta). *Botanica acta* 111: 481–489.

Rumsey FJ, Vogel JC, Russell SJ, Barrett JA, Gibby M. 1999. Population structure and conservation biology of the endangered fern *Trichomanes speciosum* Willd.(Hymenophyllaceae) at its northern distributional limit. *Biological Journal of the Linnean Society* 66: 333–344.

Rumsey FJ, Barrett JA, Gibby M, Russell SJ, Vogel JC. 2005. Reproductive strategies and population structure in the endangered pteridophyte *Trichomanes speciosum* (Hymenophyllaceae: Pteridophyta). *Fern Gazette* 17: 205–215.

Sanz-Azkue I, Díez-López J, Olariaga-Ibarguren I. 2013. Inventory and mapping of red-listed vascular flora in Hernani municipality (Gipuzkoa, Basque Country). *Munibe Ciencias Naturales* 61: 7–32.

Sanz-Azkue I, Olariaga I, Felipe A. 2014. *Vandenboschia speciosa* (Willd.) G. Kunkel en Jaizkibel: nuevas poblaciones con alta representación de su fase gametofítica y modelización de su distribución. *Munibe Monographs. Nature Series* 2: 91–97.

Schneller JJ, Holderegger R. 1996. Genetic variation in small, isolated fern populations. *Journal of Vegetation Science* 7: 113–120.

Suárez-Santiago VN, Blanca-López G, Romero-García A. 2010. *Consultoría y asistencia técnica para el estudio genético de cuatro especies del grupo de pteridófitos en peligro: Culcita macrocarpa, Pteris incompleta, Diplazium caudatum, Vandenboschia speciosa*. Sevilla: Consejería de Medio Ambiente, Junta de Andalucía.

Suter M, Schneller JJ, Vogel JC. 2000. Investigations into the genetic variation, population structure, and breeding systems of the fern *Asplenium trichomanes* subsp. *quadrivalens*. *International Journal of Plant Sciences* 161: 233–244.

Svenning JC. 2003. Deterministic Pliocene-Pleistocene extinctions in the European cool-temperate tree flora. *Ecology Letters* 6: 646–653.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358–1370.

Supporting Information

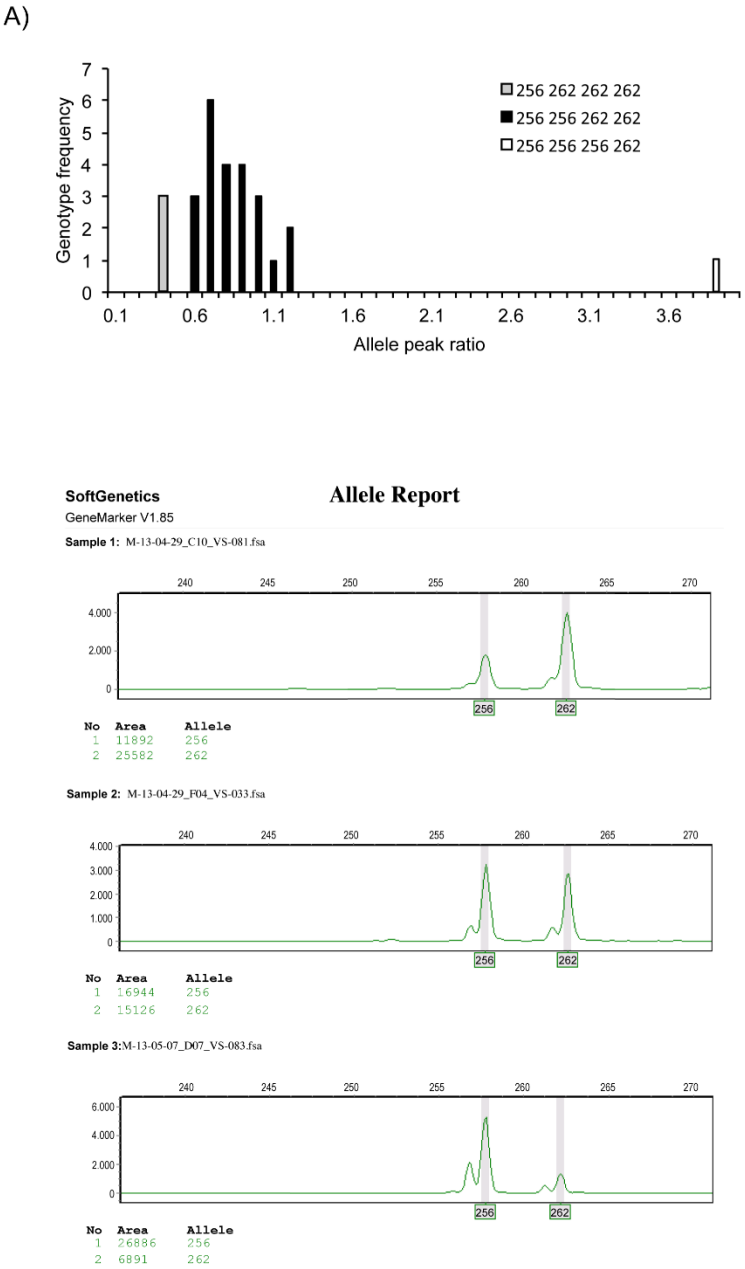
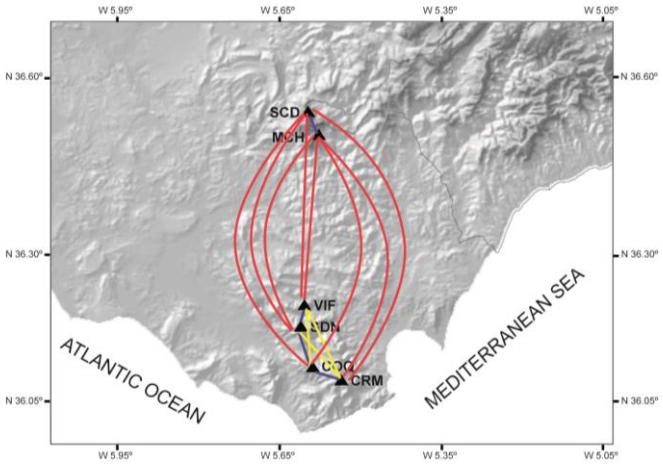


Figure S1. Example of the application of the MAC-PR method (Esselink *et al*., 2004), to estimate the allele-dosage in the heterozygous individuals of *V. speciosa*, illustrated with one allele pairwise combination for loci VS-CONS12 (A) and VS-GA87 (B) respectively. The histograms of the ratios of the allele-peak areas produce well-separated groups that correspond with different genotypes. Electropherograms exemplifying the difference in allele-dosage are also shown.



	CRM	SCD	VIF	SDN	COQ	MCH
CRM	0	3	2	2	1	3
SCD	3	0	3	3	3	1
VIF	2	3	0	1	2	3
SDN	2	3	1	0	1	3
COQ	1	3	2	1	0	3
MCH	3	1	3	3	3	0

Figure S2. Map and matrix for the geographic distance classes used in the spatial autocorrelation analysis. Short distances: distance between neighbouring populations; ≤ 6,850 m, blue lines on the map and 1 in the matrix. Medium distances: distances between populations with one or two populations in between; 6,850 – 15,000 m, yellow lines on the map and 2 in the matrix. Long distances: > 15,000 m, red lines on the map and 3 in the matrix.

Chapter 2. Phylogeographical analyses of a relict fern of palaeotropical geoflora (*Vandenboschia speciosa*): distribution and diversity model in relation to the geological and climate events of the late Miocene and early Pliocene

Samira Ben-Menni Schuler, Hammadi Hamza, Gabriel Blanca, Ana Teresa Romero-García and Víctor N. Suárez-Santiago

Manuscript in preparation

Abstract

-Fern phylogeography can be considered to be in its infancy. Most phylogeographical studies examine how Pleistocene glaciations influenced angiosperm biogeographical patterns, but not distribution and diversity patterns or their underlying processes, of pre-Pleistocene ferns. Here we reconstruct the evolutionary history of *Vandenboschia speciosa*, hypothesised to be of Tertiary palaeotropical geoflora with a peculiar perennial gametophyte.

-We sequenced 40 populations across the species range in one plastid region and two paralogues of the nuclear *gapCp* gene and conducted time-calibrated phylogenetic, phylogeographical, and species-distribution modelling analyses.

-*Vandenboschia speciosa* had a Tertiary origin. Late Miocene aridification possibly caused the long persistence in independent refugia on Eurosiberian Atlantic and Mediterranean coasts, with the independent evolution of gene pools resulting in two evolutionary units. The Cantabrian Cornice, a major refuge, could also be a secondary contact zone during Quaternary glacial cycles. Central European populations resulted from post-glacial multiple long-distance dispersals. *Vandenboschia speciosa* reached Macaronesia during the Pliocene-Pleistocene, with a phylogeographical link between the Canary Islands, Madeira, and southern Iberia, and between the Azores and north-western Europe.

-Our results support the idea that the geological and climate events of the late Miocene/early Pliocene shifted Tertiary fern distribution patterns in Europe.

Key words: Fern phylogeography, *gapCp* gene, plastid DNA, Pleistocene, refugia, species distribution modelling, Tertiary, *Vandenboschia speciosa*

1. Introduction

Palaeotropical geoflora, predominantly evergreen and thermophilous plants, populated the Northern Hemisphere, occupying a mid-latitude climate belt on both coasts of the Tethys Sea from the Late Cretaceous to the Late Miocene (Mai, 1989, 1991). In Europe, where the temperature and humidity conditions during the Paleogene were typical of a temperate subtropical zone, lauroid-type flora originated and developed (Barrón, 2003; Rodríguez-Sánchez & Arroyo, 2008). In these Tertiary lauroid forest the herbaceous layer was composed mainly of ferns (Pichi-Sermolli, 1979, 1991; Barrón, 2003; Barrón & Peyrot, 2006). The gradual (sub)tropical climate deterioration during Tertiary, especially from the mid-Miocene onwards, together with Pleistocene glaciations, caused the decline of this geoflora, which contracted to the south and west of Europe, where the climate remained suitable (e.g. Hewitt, 1996, 2003; Vogel *et al*., 1999a; Barrón, 2003; Svenning, 2003; Svenning & Skov, 2007; Ivanov *et al*., 2011). Moreover, during the Pleistocene glacial maxima, many temperate species not only found refugia at lower latitudes but also persisted in refuges at higher latitudes close to, or even within, the limits of the ice sheets (Stewart & Lister, 2001; Provan & Bennett, 2008). From all these refugia, temperate and palaeotropical species expanded their populations, even recolonizing Europe, during post-glacial periods (Taberlet *et al*., 1998; Hewitt, 1999; Stewart & Lister, 2001).

The dynamics of population contraction and/or expansion, forced by historical geological and climatic events, have influenced the distribution and diversity patterns of the species, leaving a genetic trace in their populations (Hewitt 1999, 2000; Avise, 2000) which can be identified by phylogeographical analyses (see Provan & Bennett, 2008). In plant phylogeography, most studies examine the effects of Pleistocene glaciations on the plant distributions (Petit *et al*., 2005), while few studies address the history of ancient taxa, such as that of the lauroid forest, or analyse the processes that shaped their distribution patterns (e.g. Rodríguez-Sánchez *et al*., 2009; García-Verdugo *et al*., 2013; Chen *et al*., 2014). Few phylogeographical studies on ferns are available (e.g. Trewick *et al.*, 2002; Hunt *et al*., 2009; Jiménez *et al*., 2010; Bystriakova *et al.*, 2014, Maccagni *et al.*, 2017) and, to date, no phylogeographical study, including a time-calibrated phylogeny, has attempted to reconstruct the biogeographical history of a presumed Tertiary fern species. Many European ferns are considered Tertiary relicts localized in shelters having similar microclimatic conditions to those of that time

(differing from the more general climate regime of the surrounding areas), as those found along the European Atlantic Coast and Macaronesia (Jermy, 1984; Cronk, 1992; Vogel *et al*., 1999b; Quintanilla *et al*., 2007; Vanderpoorten *et al*., 2007; Calleja *et al*., 2009; Liu & Schneider, 2013). The distribution model of these ferns, the antiquity of their lineages, and the biological peculiarities that determine their colonization dynamics, make them suitable species to explore the responses of European plants to the past climate changes and the genetic fingerprint left by these changes.

Vandenboschia speciosa (the Killarney fern; Hymenophyllaceae) is a Macaronesian-European endemic fern, which is thought to have belonged to the Tertiary laurel forests (Barrón & Peyrot, 2006). It is the only European representative of a predominantly tropical genus (Rumsey *et al.*, 1998a). One of the most remarkable features of this species is that both phases of the life cycle (sporophyte and gametophyte) are perennial and the fern can propagate vegetatively (making it a unique fern in Europe). Therefore, mixed and independent sporophyte and gametophyte populations occur throughout its distribution (Rumsey *et al.*, 1998b). Moreover, the ecological requirements of both phases vary, so the distribution of the populations depends on the generation. The sporophyte is sensitive to strong luminosity and needs high environmental moisture (e.g. shady ravines of rivers); while the gametophytes tolerate a wider range of climatic conditions, with different intensities of light and humidity (drier and darker areas such as caves or caverns), conditions under which they can develop and reproduce in a vegetative way without the need of a sporophyte (Rumsey *et al*., 1999). Thus, both the sporophyte and gametophyte appear disjunctly in the European Atlantic coast and the Macaronesian islands (Azores, Canaries, and Madeira), zones considered Tertiary flora refuges where relict populations could survive the glacial periods. In these areas, both generations could develop a normal fern life cycle (Rumsey *et al*., 2005). As the species entered the continent or expanded in latitude (to Northern and Eastern Europe, reaching Scotland and Poland, respectively), the sporophyte generation disappeared and only the gametophyte remained.

Thus far, only an approximation has been made to the phylogeographical study of *V. speciosa*, using PCR-RFLPs (Rumsey & Sheffield, 1996), and a few studies geographically limited on genetic diversity using allozymes (Rumsey *et al*., 1998b, 1999, 2005) and microsatellites (Ben-Menni Schuler *et al*., 2017). Despite these studies, the origin and evolutionary history of *V. speciosa* in Europe and Macaronesia remains

elusive. In the present study, we used plastid DNA (cpDNA; intergenic spacer *trn*H-*psb*A) and low-copy nuclear gene (*gapCp* gene) sequences, and species distribution modelling (SDM) to explore patterns of diversity and distribution of *V. speciosa* throughout Europe and Macaronesia. Within our overall aim, the specific objectives were: (1) to test the Tertiary relict hypothesis; (2) to compare the genetic structure between the two life-cycle phases; (3) to infer the phylogeographical patterns throughout its whole range and explore the putative impacts of past climate changes in modelling those patterns; and (4) to determine the impact of future climate warming on the species distribution as a whole, considering a possible differential response of gametophyte and sporophyte.

2. Materials and Methods

2.1. Plant material

Samples of *Vandenboschia speciosa* (Willd.) G. Kunkel (synonym *Trichomanes speciosum* Willd.) were taken from 40 populations in 11 geographical regions across the fern's distribution range (Fig. 1; Supporting Information Table S1). The number of populations per region ranged between two and seven. Twelve populations were formed only by gametophytes, four by only sporophytes, and the remaining were mixed populations. Both gametophyte and sporophyte generations were sampled, including between one and 11 individuals per generation and population. A total of 309 individuals were sampled, i.e. 168 gametophytes and 141 sporophytes. Fronds of *V. davallioides*, and two *V. birmanica* accessions were also sampled, from herbaria, as outgroup species (Table S1).

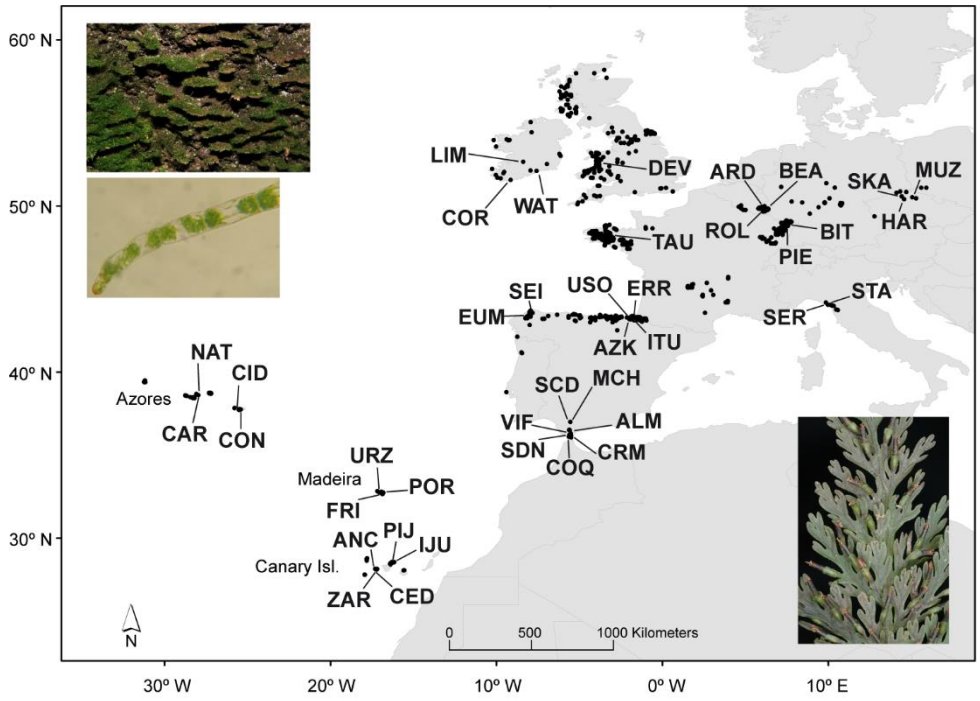


Fig. 1 Geographical distribution of *Vandenboschia speciosa*, showing the 40 populations sampled for phylogeographical analyses (see Table S1 for population code) and the location of presence records (black dots) used for species distribution modelling (SDM). Two photographs of the gametophyte, showing the habit and a microscopic detail, are shown in the upper left corner; and in the lower right corner the detail of a frond with sori is shown.

2.2. DNA extraction, PCR amplification, and sequencing

Total genomic DNA of the 312 sampled individuals was extracted from cleaned filaments of gametophytes, and from fresh, silica dried, or herbaria fronds of sporophytes using the NucleoSpin Plant II kit (Macherey-Nagel GmbH & Co. KG; Düren, Germany), following the manufacturer's instructions.

Plastid DNA for the intergenic spacer *trn*H-*psb*A was amplified by polymerase chain reaction (PCR) in all 309 individuals of *V. speciosa* and two outgroup species; however, this region could not be amplified in one of the two *V. birmanica* accessions. We used the *gapCp* gene as nuclear marker, which has been used successfully in various phylogenetic studies of Hymenophyllaceae species, including the genus *Vandenboschia*, and has shown variation at infraspecific levels (e.g. Ebihara *et al*., 2005, 2009; Nitta *et al*., 2011). Due to the tetraploid nature of *V. speciosa* (Manton, 1950; although recent flow cytometry data from Irish populations suggest that there could be a mix of cytotypes; Ní Dhúill, unpubl. Ph.D), only gametophytes (diploids) were used for nuclear marker amplification. In total, 150 *V. speciosa* gametophytes

from 36 populations and one sporophyte of each *V. birmanica* accessions, and *V. davallioides* were analysed for the nuclear *gapCp* gene. See Methods S1 for details of PCR conditions.

For both cpDNA and nuclear *gapCp,* sequences were edited and aligned, using the Clustal algorithm, in the alignment editor BIOEDIT v7.0.5.3 (Hall, 1999), and then adjusted by eye. All sequences were deposited in the European Nucleotide Archive (accessions XXXXX–XXXXX).

2.3. Identification of *gapCp* paralogues by phylogenetic analysis

Visually inspecting the *gapCp* alignment, we detected a set of nucleotide positions that differentiate two types of sequences. To establish the homology relationship between these variants, we made a maximum likelihood (ML) phylogenetic analysis with PhyML v3.0 (Guindon *et al*., 2010), including our *V. speciosa*, *V. birmanica*, and *V. davallioi*des sequences, and the *gapCp* sequences in GenBank for *V. radicans* group and *V. auriculata* (subgenus *Lacosteopsis*; as outgroup species) (Ebihara *et al*., 2005; accession numbers: AB196370-AB196419). The data matrix included a total of 499 sequences after we removed the inter-copy recombinant and single-point mutation sequences, detected with DNAsp v5.10 (Librado & Rozas, 2009), and we used haplotypes instead all sequences (see Methods S1 for details).

Because the two *gapCp* variants were paralogues (called gapCp-572 and gapCp-575, see Results), we independently considered both copies in the subsequent analyses and only the largest intra-copy nonrecombining portion (detected with DNAsp).

2.4. Genetic diversity and structure

Population and regional genetic diversities were assessed by the number of haplotypes (*ha*), haplotype diversity (*Hd*), and nucleotide diversity (π) calculated with ARLEQUIN v3.5.2.2 (Excoffier & Lischer, 2010) for both the cpDNA and the two copies of *gapCp*. For cpDNA and mixed populations (gametophytes and sporophytes), diversity indices were calculated considering all individuals together (including both life-cycle phases) and also for the independent phases separately.

The distribution of genetic variability between generations (gametophyte and sporophyte) was evaluated for cpDNA using an analysis of molecular variance (AMOVA; Excoffier *et al*., 1992) and tested with a permutation test (10,000 permutations) with ARLEQUIN. Hierarchical AMOVAs were also conducted to quantify the proportion of total genetic variance explained by the difference between the 11 geographical regions and between populations within regions, for all molecular markers. In addition, two more AMOVAs were performed for cpDNA, considering the geographical distribution of the haplotypes and the results of the network analysis (see Results). Thus, we considered a highest level of population grouping called 'evolutionary unit'. The AMOVAs were made to test the differentiation between evolutionary units, and they included, first, two evolutionary units (Northern and Southern) and the Cantabrian region; and later without taking into account the Cantabrian region. Cantabrian region was considered an independent unit due to the high admixture of haplotypes from the Northern and Southern evolutionary units.

Haplotype networks were reconstructed for all molecular markers (cpDNA, gapCp-572, and gapCp-575) following the statistical parsimony method (Templeton *et al*., 1992) as implemented in TCS v1.21 (Clement *et al*., 2000).

2.5. Haplotype phylogeny and dating

Phylogenetic relationships among cpDNA haplotypes of *V. speciosa* and the outgroup species were inferred using Bayesian Inference (BI), with MrBayes v3.1.2 (Ronquist & Huelsenbeck, 2003; see Methods S1 for details).

To relate genetic differentiation found among cpDNA haplotypes to Neogene-Quaternary events, we estimated divergence times using BEAST2 package (Bouckaert *et al*., 2014; see Methods S1 for details). We followed a two-step strategy for tree calibration, because we could not establish specific dates for the few known fossils (the oldest considered to be from the Upper Miocene; see Bozukov, 2008). An initial analysis to estimate the divergence time of *V. speciosa* lineage, using an uncorrelated lognormal relaxed-clock model (Drummond *et al*., 2006), was implemented with sequences of the *rbc*L gene for the main lineages of Trichomanoids (including 15 species of *Vandenboschia*), all taken from the GenBank database (accession numbers on the resulting tree; see Results). Phylogeny was calibrated employing dates from

Schuettpelz and Pryer (2009). The second analysis was focused on the *trn*H-*psb*A haplotypes of *V. speciosa*, using the sequences of eight species of *Vandenboschia* (six sequences taken from Genbank, accession numbers on the tree in Results, and two generated by us for *V. birmanica* and *V. davallioides*) as an outgroup and the Coalescent Constant Population Model as a tree prior. Two calibration points were defined, using the dates from the first analysis.

2.6. Demographic Analyses

Neutrality tests were applied, Fu's *F* (Fu, 1997) and Tajima's *D* (Tajima, 1989), to detect possible historical demographic processes (expansion or contraction), using ARLEQUIN. Both tests were performed considering geographical regions and evolutionary units, and with each molecular marker. The level of significance of both statistics was based on 10,000 simulated samples. In addition, cpDNA sequences were used to test for evidence of population size fluctuations within *V. speciosa* evolutionary units and also within Cantabrian range by constructing Bayesian Skyline Plots with BEAST (BSP; Drummond *et al*., 2005; see Methods S1 for details).

To reconstruct the historical migration routes of *V. speciosa* and locations of internal nodes of the haplotype phylogeny, we analysed cpDNA sequences using the marginal approximation of the structured coalescent implemented in BEAST2 package MASCOT v1.2.2 (Marginal Approximation of the Structured COalescenT; Müller *et al*., 2018). We defined three areas based in the evolutionary unit structuring: North, South, and Cantabrian (see Methods S1 for details).

2.7. Species distribution modelling

To identify past refugia and future distribution areas for *V. speciosa*, we evaluated the potential range of the species, considering the two phases of the life cycle both together and separately, under past, current, and future conditions with species distribution modelling (SDM). For environmental data, we used 19 BIOCLIM variables at a resolution of 2.5 arc-minutes (*c*. 5 km). Past and current climate data were available from the WorldClim database (www.worldclim.org; Hijmans *et al*., 2005) and included data for the current-day period (1,950– 2,000), the Last Glacial Maximum (LGM; *c*.

21 ka) simulated by CCSM model (the Community Climate System Model), and for the Last Interglacial period (LIG; *c*. 120 ka). In addition, we made predictions for future climatic conditions in 2,080 for the most impacting IPCC's climate scenario: RCP8.5 (Van Vuuren *et al*., 2011) available through the CCAFS Climate portal (www.ccafs-climate.org). Highly correlated variables (Pearson's $R \geq 0.8$) were reduced to eight uncorrelated variables (Table S2) used as predictors to calibrate the distribution models. Species occurrence data is a collection of references in databases (the Global Biodiversity Information Facility data portal (http://www.gbif.org/), the Biodiversity databank of the Canary Islands

(http://www.biodiversidadcanarias.es/atlantis/common/index.jsf), and the Azores Biodiversity databank (http://www.atlantis.angra.uac.pt/atlantis/common/index.jsf), literature (Louis-Arsène, 1953; Rumsey *et al*., 1998a; Marchetti, 2002; Krukowski & Świerkosz, 2004; Loriot & Geslin, 2005; Eichler & Kempf, 2012; Sánchez-Velázquez, 2013; Sanz-Azcue *et al*., 2013; Dhúill *et al*., 2015), and our own field records. A total of 1,548 presence records (1,066 for gametophyte and 482 for sporophyte) of *V. speciosa* were finally compiled (Fig. 1). To perform the SDM, we applied Maximum Entropy Modelling implemented in the software package MAXENT 3.4.1 (Phillips *et al*., 2006). Models were generated using cross-validation of 5 replicate runs. Model performance was assessed on the basis of the area under the receiver operating characteristic curve (AUC). The contribution of each predictor variable in the model was analysed by the permutation importance and percent contribution coefficients (Table S2). A final reduced model including the most important variables (Martínez *et al*., 2012), Mean Diurnal Range and Minimum Temperature of Coldest Month, was finally computed.

3. Results

3.1. CpDNA and nuclear marker characteristics

Plastid-DNA alignment included 317 sequences in total (309 of *Vandenboschia speciosa* and eight of outgroups). In *V. speciosa trn*H-*psb*A sequences were 450 base pairs (bp) in length and eight nucleotide positions were polymorphic sites among them. Alignment for the nuclear *gapCp* comprised 665 sequences (592 of *V. speciosa* and 73 of outgroups) and it was 603 bp in length. In *V. speciosa*, we detected two types of *gapCp* sequences, which we called gapCp-572 and gapCp-575 (572 bp and 575 bp in

length, respectively). These differed by 20 inter-copy variable positions (but fixed within each copy) and five indels. We detected 32 inter-copy recombinant and 114 single point mutation sequences and deleted them from alignment for the following analyses. The number of polymorphic positions in the final alignment was 344 and the number of indels 17. ML phylogenetic analysis showed the paralogy relationships between both *gapCp* copies (Fig. S1), since gapCp-575 was closely related to the Ebihara's type-C copy (*V. birmanica*) and gapCp-572 was related to the Ebihara's type-A and type-B copies (*V. kalamocarpa* and *V. nipponica*, respectively). We also detected both *gapCp* copies in our material of *V. birmanica* and *V. davallioides*; sequences from these species were interspersed with the sequences of *V. speciosa* in the resulting phylogenetic analysis (Fig. S1). Independent alignments for *gapCp* copies, including only the largest intra-copy nonrecombining portion, were 80 bp and 92 bp in length and included 23 and 35 polymorphic sites (gapCp-572 and gapCp-575, respectively).

3.2. Genetic diversity and structure

For cpDNA, gapCp-572, and gapCp-575, the total number of haplotypes found were nine, 28 and 43, respectively. The results for the diversity indices are shown in Table 1. At the population level, the mean diversity values for the cpDNA were $Hd = 0.695$ and $\pi = 0.00312$, while for gapCp-572 were $Hd = 0.273$ and $\pi = 0.00368$, and $Hd = 0.323$ and $\pi = 0.00413$ for gapCp-575. At the geographical region level, the most diverse regions were: the Basque Country for the cpDNA; Vosges du Nord, and Italy for gapCp-572; and Italy for gapCp-575 (Table 1).

Table 1 Information on number of haplotypes and diversity indices for *trn*H-*psb*A, gapCp-572, and gapCp-575 sequences in each population or region of *Vandenboschia speciosa* analysed in the present study

	*psb*A-*trn*H				gapC-572				gapC-575			
Code	*ha*	*Priv*	*Hd*	π	*ha*	*Priv*	*Hd*	π	*ha*	*Priv*	*Hd*	π
Andalusia	2		0.425	0.0019	4	1	0.193	0.0087	8	4	0.340	0.0045
ALM	1		0	0	2		0.285	0.0124	2	1	0.182	0.0019
COQ	1		0	0	2	1	0.222	0.0096	3	1	0.416	0.0048
CRM	1		0	0								
MCH	2		0.355	0.0016	1		0	0	4	1	0.371	0.0058
SCD	1		0	0								
SDN	1		0	0								
VIF	1		0	0	2		0.250	0.0108	3	1	0.464	0.0054
Azores	4	2	0.233	0.0008	7	2	0.326	0.0153	8	3	0.371	0.0048
CAR	3	2	0.377	0.0009	2		0.200	0.0087	3	1	0.345	0.0055
CID	2		0.200	0.0009	3		0.524	0.0248	3	1	0.345	0.0039
CON	2		0.200	0.0009	3	1	0.464	0.0217	3	1	0.318	0.0036
NAT	2		0.200	0.0009	2	1	0.222	0.0097	3		0.472	0.0055
Basque Country	5	1	0.673	0.0030	6	1	0.331	0.0155	4	3	0.221	0.0025
AZK	3		0.644	0.0029	4	1	0.643	0.0326	2	1	0.286	0.0031
ERR	1		0	0	1		0	0	1		0	0
ITU	1		0	0	1		0	0	1		0	0
USO	5	1	0.767	0.0036	3		0.378	0.0174	3	2	0.464	0.0054
Canary Isl.	3	1	0.405	0.0013	5	3	0.192	0.0087	8	4	0.348	0.0041
ANC	1		0	0	1		0	0	2		0.333	0.0036
CED	1		0	0	2	1	0.286	0.0124	3	1	0.524	0.0062
IJU	3		0.711	0.0032	3	1	0.417	0.0193	2		0.222	0.0024
PIJ	2		0.467	0.0010	2	1	0.182	0.0079	3	1	0.345	0.0040
ZAR	2		0.467	0.0010	1		0	0	4	2	0.396	0.0047
Czech Republic	1		0	0	3	1	0.257	0.0116	6	5	0.447	0.0076
HAR	1		0	0	2	1	0.400	0.0174	1		0	0
MUZ	1		0	0	2		0.667	0.0290	5	4	0.786	0.0163
SKA	1		0	0	1		0	0	2	1	0.333	0.0036
Galicia	3		0.567	0.0025	4	1	0.249	0.0113	6	3	0.490	0.0060
EUM	2		0.533	0.0024	2		0.133	0.0058	6	3	0.778	0.0109
SEI	2		0.356	0.0016	3	1	0.464	0.0218	1		0	0
Ire-Wal-Bri*	4	1	0.611	0.0026	5	3	0.197	0.0089	10	7	0.379	0.0051
COR	2	1	0.200	0.0004	3	1	0.295	0.0134	2	1	0.222	0.0024

DEV	1	0	0	1		0	0	1		0	0
LIM	1	0	0	1		0	0	5	2	0.576	0.0072
TAU	2	0.533	0.0024	2	1	0.667	0.0290	1		0	0
WAT	2	0.467	0.0021	2	1	0.222	0.0097	5	4	0.667	0.0109
Italy	1	0	0	7	3	0.521	0.0261	6	4	0.515	0.0077
SER	1	0	0	5	2	0.667	0.0348	3	2	0.417	0.0072
STA	1	0	0	3	1	0.378	0.0174	4	2	0.643	0.0081
Luxembourg	1	0	0	4	2	0.331	0.0153	2	1	0.100	0.0011
ARD	1	0	0	3	1	0.524	0.0248	1		0	0
BEA	1	0	0	1		0	0	1		0	0
ROL	1	0	0	2	1	0.333	0.0145	2	1	0.333	0.0036
Madeira	2	0.351	0.0015	2		0.133	0.0058	1		0	0
FRI	2	0.467	0.0021	2		0.200	0.0087	1		0	0
POR	1	0	0	1		0	0	1		0	0
URZ	2	0.400	0.0018								
Vosges du Nord	2	0.356	0.0016	4	1	0.533	0.0261	4		0.350	0.0041
BIT	2	0.600	0.0027	2		0.400	0.0174	1		0	0
PIE	1	0	0	3	1	0.700	0.0348	4		0.714	0.0093

*Ireland-Wales-Brittany; *ha*, number of haplotypes; *Hd*, haplotype diversity; π, nucleotide diversity; *Priv*, private haplotype.

The representation on a map of the cpDNA haplotype frequencies and distributions suggests a geographical structuring (Fig. 2a). Haplotypes H-II and H-III are the most frequent and widespread, showing a mainly differentiated distribution. Haplotype H-II is dominant mainly into the southern geographical regions (Andalusia, Canary Islands, Madeira, and Italy) but also appears in the northern Iberian Peninsula (Basque Country, Galicia) as well as in Brittany, Ireland, and Wales. On the other hand, H-III dominates in the central and northern geographical regions (populations from Azores, Brittany, Czech Republic, Ireland, Luxemburg, Vosges du Nord) and coexists with H-II in northern Iberian (Basque Country and Galicia). Other haplotypes distributed mainly in the southern regions are H-I, in all southern regions except Italy and scantily represented in northern Iberian and Luxemburg, whereas H-VIII private to the Canary Islands. Finally, other central and northern haplotypes are: H-VI, in Azores, Basque Country, Czech Republic (the only one found), and Ireland; and the private and single-individual haplotypes H-IV and H-V (Azores), H-VII (Ireland), and H-IX (Basque Country). Plastid-DNA network analysis resulted in the network shown in

Figure 2b. According to the geographical distribution of cpDNA haplotypes and their relationships in the network, we defined two supra-regional population groupings that we termed “evolutionary units” (North: ARD, BEA, BIT, CAR, CID, CON, COR, DEV, HAR, LIM, MUZ, NAT, PIE, ROL, SKA, TAU, WAT and South: ALM, ANC, CED, COQ, CRM, IJU, FRI, MCH, PIJ, POR, SCD, SDN, SER, STA, URZ, VIF, ZAR). Populations in the Cantabrian region were considered as belonging to an independent region due to the high admixture of haplotypes from the Northern and Southern evolutionary units.

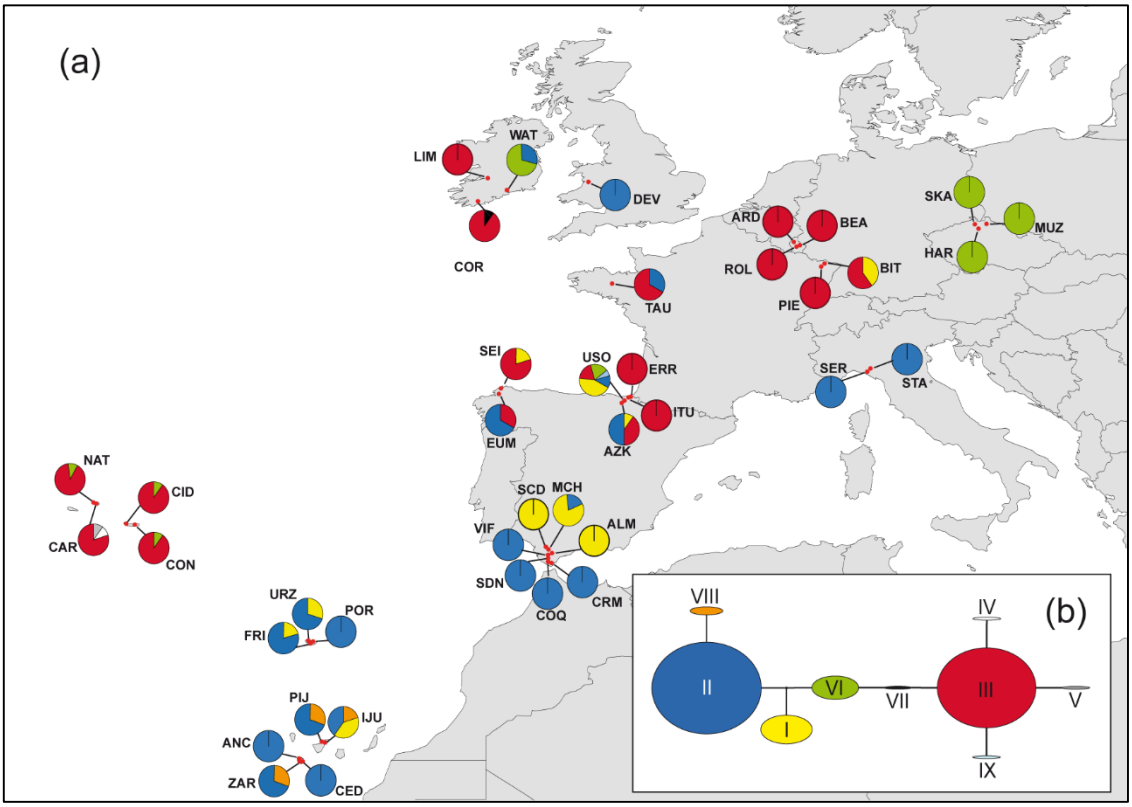


Fig. 2 Distribution and network of the cpDNA haplotypes detected in the populations sampled of *Vandenboschia speciosa*. (a) Geographical distribution of the cpDNA haplotypes among the 40 populations sampled (see Table S1 for population code); pie charts indicate haplotype frequency. (b) Inferred cpDNA network, following the statistical parsimony method, with TCS; haplotypes are denoted by Roman numerals, and circle sizes are proportional to the haplotype frequencies.

For *gapCp*, each copy has a main haplotype which is present along the 40 sampled populations. Of the 27 and 42 minority haplotypes found for gapCp-572 and gapCp-575, respectively, most were single-individual haplotypes. Network analyses resulted in only one star-like-shaped haplogroup for each *gapCp* copy (Fig. S2). No *gapCp* copy showed apparent haplotype geographical structuring (data not shown).

AMOVA analyses of cpDNA sequences showed no genetic differentiation between sporophyte and gametophyte (F_{CT} =0.011, P = 0.215), indicating that the genetic variation resides mainly among populations (68.95%, $P < 0.001$) rather than between generations (Table 2). This result induced us to combine gametophyte and sporophyte data in all remaining analyses, and not to consider the two phases separately. When the 11 geographical regions were considered, almost 39% ($P < 0.001$) of variation was between regions (Table 2). Considering supra-regional groupings, only when we excluded the Cantabrian region did a clear differentiation between Northern and Southern units become evident (F_{CT} = 0.51, 50.77% of variation, $P < 0.001$; Table 2).

AMOVA analyses of *gapCp* sequences and 11 regions revealed that almost all the genetic variation in both gapCp-572 and gapCp-575 copies resides within populations (99.63% and 99.13%, respectively; Table 2).

Table 2 Hierarchical analysis of molecular variance (AMOVA)

Source of variation	d.f	Sum of squares	Variance components	Percentage of variation	Fixation indices	*P-value**
*trn*H-*psb*A						
Among generations	1	4.042	0.00773	1.084	F_{CT} = 0.011	0.21495
Among populations within generations	62	160.115	0.49222	68.954	F_{SC} = 0.7	**<.001**
Within populations	245	52.400	0.21388	29.962	F_{ST} = 0.7	**<.001**
Total	308	216.557	0.71383			
Among geographical regions	10	48.120	0.14279 Va	38.94	F_{CT} = 0.39	**<.001**
Among populations within regions	29	25.525	0.09971 Vb	27.19	F_{SC} = 0.45	**<.001**
Within populations	269	33.417	0.12423 Vc	33.87	F_{ST} = 0.66	**<.001**
Total	308	107.061	0.36672			
Among evolutionary units (3 units)	2	34.062	0.16218 Va	39.47	F_{CT} = 0.39	**<.001**
Among populations within units	37	39.583	0.12447 Vb	30.29	F_{SC} = 0.5	**<.001**
Within populations	269	33.417	0.12423 Vc	30.23	F_{ST} = 0.7	**<.001**
Total	308	107.061	0.41087			
Among evolutionary units (2 units)	1	31.022	0.23645 Va	50.77	F_{CT} = 0.51	**<.001**
Among populations within units	32	33.948	0.13004 Vb	27.92	F_{SC} = 0.57	**<.001**
Within populations	220	21.833	0.09924 Vc	21.31	F_{ST} = 0.78	**<.001**
Total	253	86.803	0.46572			
gapCp-572 copy						
Among geographical regions	10	1.298	0.00037 Va	0.27	F_{CT} = 0.003	0.34851
Among populations within groups	25	3.066	-0.00218 Vb	0	F_{SC} = -0.016	0.72851
Within populations	235	32.544	0.13848 Vc	99.73	F_{ST} = -0.013	0.76386
Total	270	36.908	0.13667			
gapCp-575 copy						
Among regions (11 groups)	10	1.651	-0.00084 Va	-0.50	F_{CT} = -0.005	0.69554
Among populations within groups	24	4.518	0.00230 Vb	1.36	F_{SC} = 0.013	0.14871
Within populations	275	46.183	0.16794 Vc	99.14	F_{ST} = 0.008	0.25485
Total	309	52.352	0.1694			

*Statistically significant values are indicated in bold text.

3.3. CpDNA haplotype phylogeny and dating

The phylogenetic tree resulting from the Bayesian analysis with MrBayes (Fig. S3), showed how the haplotypes of *V. speciosa* formed a monophyletic group (posterior probability, pp = 1). Haplotype relationships agree with those found in the cpDNA network: H-III, H-IV, H-V, H-VII, and H-IX form a clade (pp = 0.82); haplotypes H-II and H-VIII group together; and relationships of haplotypes H-I and H-VI remain ambiguous.

The first step in the divergence-time estimates resulted in the beginning of the diversification of *V. davalloides* group 20.53 Ma (95% HPD: 10.26-31.95), during the transition between the Aquitanian and the Burdigalian from the Miocene, and that of the *V. nipponica-kalamocarpa* group 12.21 Ma (95% HPD: 5.69-19.77) during the Serravallian (middle Miocene) (nodes A and B in Fig. S4). Using these datings as secondary calibration points, we dated the *trn*H-*psb*A haplotype nodes of *V. speciosa* (Fig. 3). The analysis suggests that *V. speciosa* lineage dated back to the beginning of the Serravallian (13 Ma, 95% HPD: 2.99-15.99), which the start of the diversification of its haplotypes occurred 5.1 Ma (transition between Miocene and Pliocene; 95% HPD: 2.41-14.29), that diversification of the haplogroup including H-III began in the Pliocene-Pleistocene transition (2.58 Ma), and that H-VIII diverged from H-II 1.27 Ma (Fig. 3).

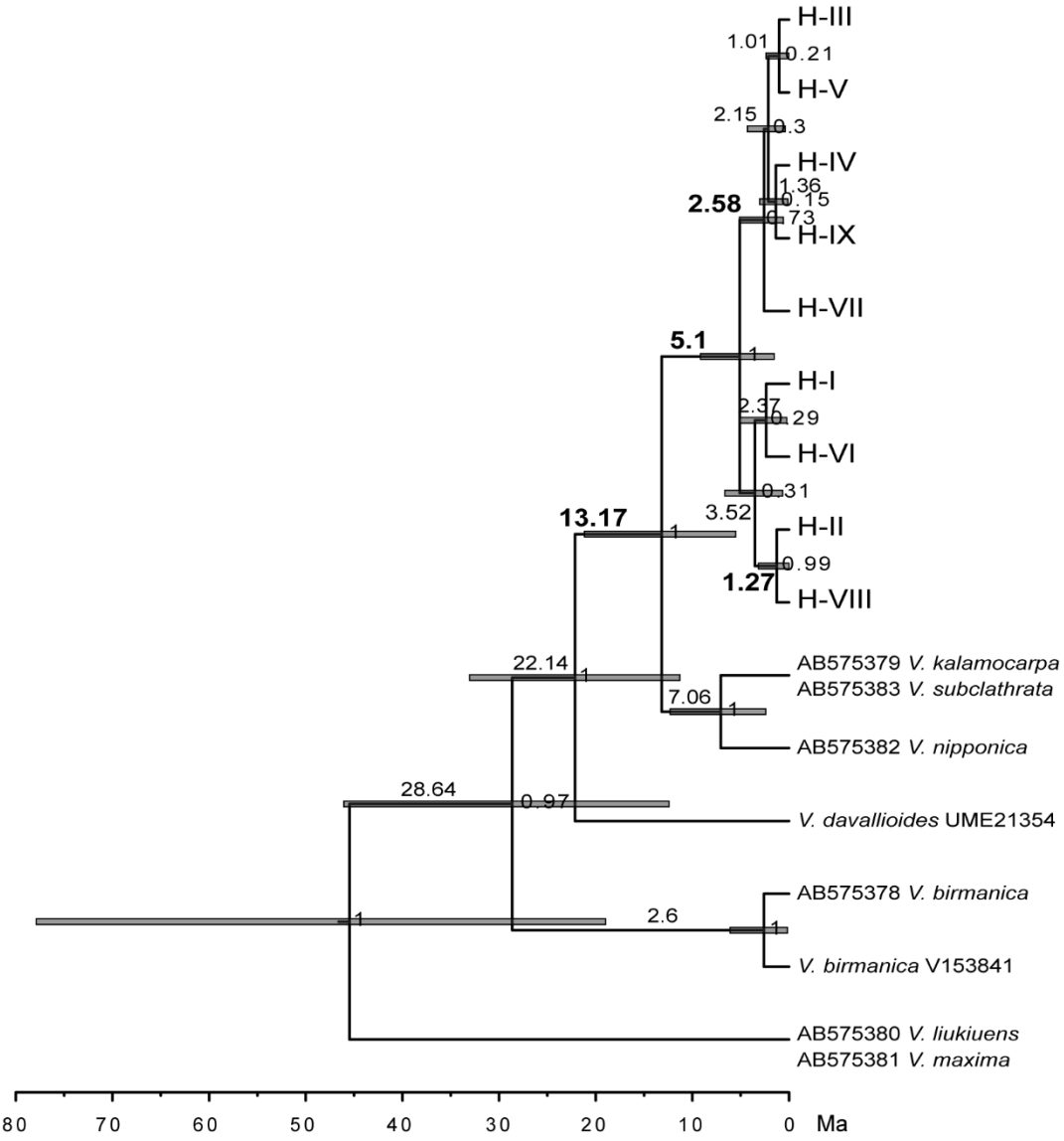


Fig. 3 Time-calibrated phylogeny for the *Vandenboschia speciosa* cpDNA haplotypes, and outgroup species, derived from BEAST. Numbers above branches are the mean divergence ages (in million years ago) for each node (the ages for the key nodes of *V. speciosa* are shown in bold); grey bars represent 95% highest posterior density intervals for each node; numbers after nodes are BEAST posterior probabilities; the time scale is printed in million years ago (Ma).

3.4. Demographic analysis of *V. speciosa*

For cpDNA, the neutrality tests (Fu, Li, and Tajima's test) turned out to be non-significant (Table S3); however, negative values close to significance were registered for the Azores region. At the evolutionary-unit level, the Northern unit was the only one that had negative values (not significant). The neutrality tests made with the nuclear markers in all cases resulted in negative and significant values for all geographical regions and evolutionary units.

The BSP analyses found evidence for range expansion only for the Northern evolutionary unit, this starting from 100 to 80 thousand years ago (Ka; beginning of the Würm glaciation) (Fig. 4a).

The structured coalescent approach identified migrations occurring mainly between the Northern and Cantabrian regions (Fig. 4b), with a greater proportion of individuals from the Cantabrian region coming from the north (mean: 1.76), although a relatively high migration rate was detected in the opposite direction (mean: 1.12). The migration rate from the south to the Cantabrian region was close to unity (mean: 0.96), while the rest of the rates proved extremely low (means < 0.6). On the other hand, MASCOT analysis yielded in the ancestral locations of *V. speciosa* lineages displayed in Figure 4c. The South was the root state with moderate probability (0.65), where haplotypes H-I and H-II originated (probabilities: 0.6 and 0.86, respectively) and from which they migrated into the Cantabrian and northern regions. The haplotype H-VIII, private to the Canaries, originated there from H-II. Haplotype H-VI originated in the North (probability: 0.82) and introduced into the Cantabrian Cornice. Finally, the origin of haplotype H-III was resolved as the North with little more than the 50% probability (0.58), and the remaining haplotypes (private to different regions) originated from *in situ* H-III.

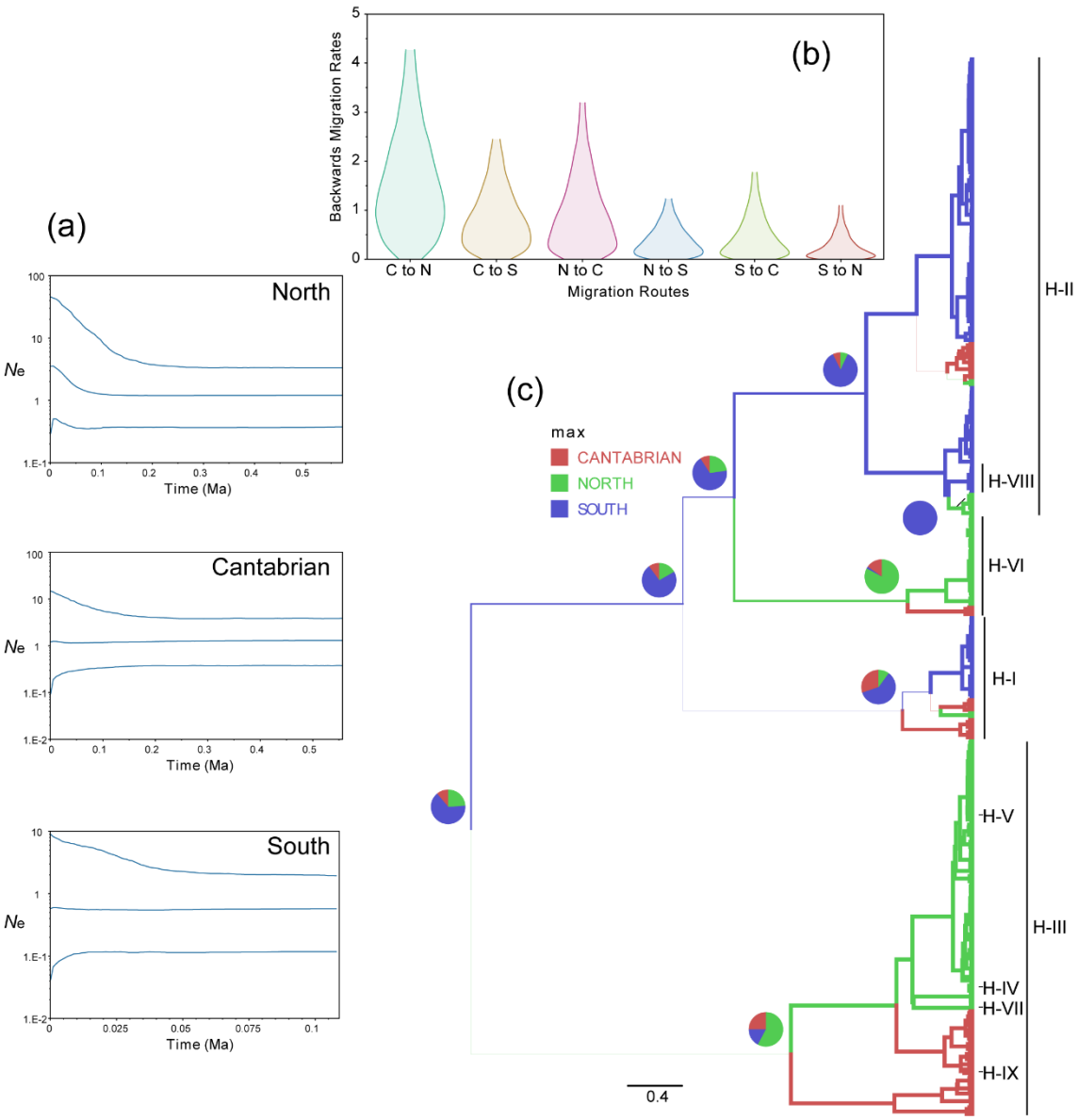


Fig. 4 Demographic analyses based on coalescence considering the supra-regional groupings of *Vandenboschia speciosa* (evolutionary units and the Cantabrian Cornice), inferred from cpDNA and implemented with BEAST. (a) Bayesian Skyline Plots depicting changes in effective population size (N_e) as a function of time (in million years ago, Ma); in each plot, the centre line is the median estimate, and the upper and lower lines delimit the highest posterior density (HPD) 95% confidence intervals for N_e. (b) Violin plot of the inferred backwards migration rates between the different regions (C, N, and S: Cantabrian Cornice, Northern and Southern evolutionary units, respectively), using the marginal approximation of the structured coalescent (MASCOT) as a population prior. (c) Maximum clade credibility tree showing the inferred root regions for the cpDNA haplotype, using MASCOT as a population prior; the pie charts show the inferred probability of the root being in any of the three regions.

3.5. Species distribution modelling

For all models, the AUC values were high (minimum value of AUC = 0.962). For the species (gametophyte and sporophyte together), the MAXENT current and LIG predictions showed regions of suitable habitats that coincided largely with the current

distribution, with additional areas of the distribution range on the European Atlantic coasts further north and more widely along the Mediterranean sea, where the species is currently absent (Fig. 5a,b). Palaeodistribution modelling for LGM suggested a strong contraction of the suitable habitats in Northern and Central Europe (Fig. 5c). According to LGM output, refugia were located in Macaronesia, the European Atlantic coast and a few regions on the Mediterranean coastline (Fig. 5c). The Atlantic coastal strip of the Iberian Peninsula, from Galicia to the south, appeared as a continuous zone of high suitability. This continuity was not found in the LIG predictions (Fig. 5b). The MAXENT future projections, using the RCP8.5 scenario, suggested a partial reduction of the suitable habitats of the species on the coasts of Portugal, northern Iberian Peninsula, and Macaronesian islands, along with an increase of the suitable habitats northwards of the European Atlantic coast (Fig. 5d). Independent analyses for gametophyte and sporophyte resulted in similar projections in both generations for the past, present, and future, except that the gametophyte presented a higher distribution than did the sporophyte, especially towards the north and the east of the continent (Fig. S5).

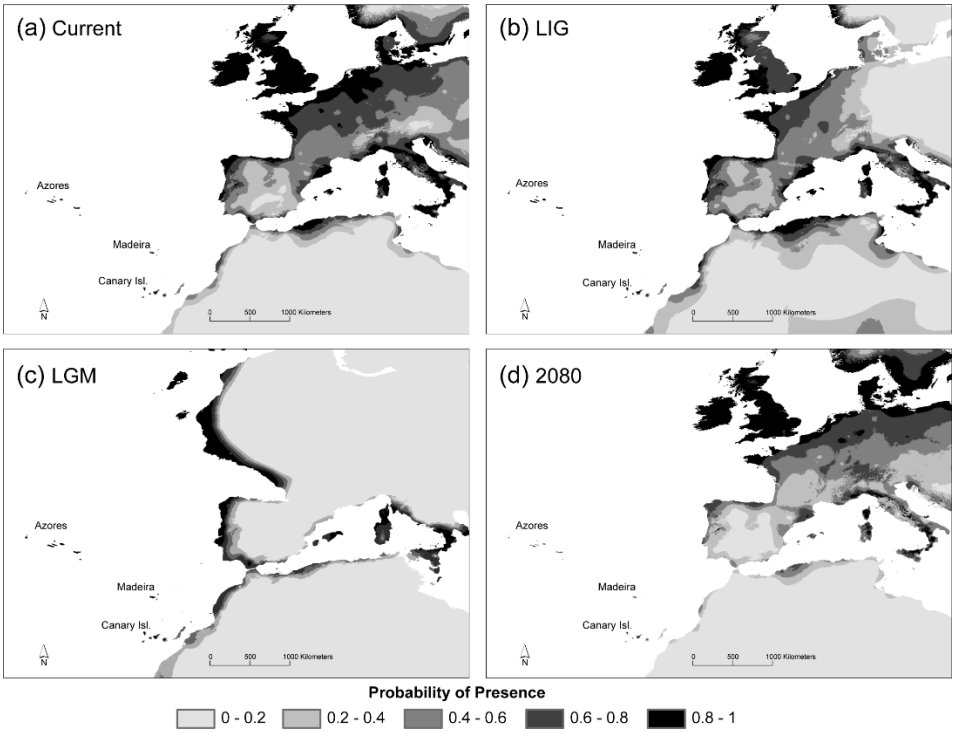


Fig. 5 Maps of predicted environmental suitability for *Vandenboschia speciosa* using the maximum entropy algorithm and the Community Climate System Model (CCSM), as implemented in MAXENT. Map showing the projections for current (a), last interglacial (LIG, c. 120-140 kya) (b), last glacial maximum (LGM, c. 21 kya) (c), and future (2080) (d) are shown for the species as a whole. The probability of presence is shown as continuous values from the threshold (defined as Maximum Sensitivity plus Specificity) to 1. The figure was generated with ARCGIS.

4. Discussion

4.1. Absence of genetic structure between the sporophytic and gametophytic phases of *V. speciosa*

Vandenboschia speciosa is one of the few fern species with a long-lived gametophytic phase (*cf*. Pinson *et al.*, 2017). Gametophytes differ from sporophytes in ecological tolerances (Makgomol & Sheffield, 2001, 2005) and dense clonal populations can result without sporophytes in continental environments, beyond the oceanic environments to which sporophytes are limited. Despite this, our AMOVA results showed no structuring of plastid genetic diversity between gametophytes and sporophytes (Table 2); as Rumsey *et al*. (1999) observed using allozyme banding phenotypes in three sites of south-western Scotland.

4.2. Tertiary origin of *V. speciosa*

Our results show a Tertiary origin for *V. speciosa* (see Notes S1 for discussion on the possible allopolyploid origin of *V. speciosa*), at the end of the middle Miocene, placing it as part of the palaeotropical geoflora that inhabited on both sides of the Tethys Sea from the Cretaceous until the end of the Miocene (Mai, 1989, 1991; Erdei *et al*., 2007; Utescher *et al*., 2007). This dating agrees well with the oldest known fossils of the species in Georgia (Upper Miocene; Bozukov, 2008). The location of these fossils and the phylogenetic position of *V. speciosa* within a clade of species from Japan and South-east Asia (Fig. S4) suggest an Asian origin of the ancestors of *V. speciosa*. A possible ancestor could be *Trichomanes sacci* (= *T.* cf. *radicans*), a *Vandenboschia* species that inhabited Eastern Europe from the Upper Oligocene to Middle Miocene (Palamarev & Petkova, 1990; Bozukov, 2008).

Currently *V. speciosa* is undergoing strong geographical isolation from its closest relatives in South-east Asia. This disconnection could have occurred during the Upper Miocene (Fig. 3, Fig. S4), when the brusque climate changes in Asia (Molnar *et al*., 2010; Miao *et al*., 2012) could have provoked the extinction of the close relatives from the areas between Europe and South-east Asia; a scenario proposed also for *Davallia canariensis* (Liu & Schneider, 2013).

4.3. Climate-change-driven phylogeographical history of *V. speciosa*

The genetic structure detected for *V. speciosa* across its distribution range, with the plastidial marker *trn*H-*psb*A (Fig. 2a), indicates the presence of two main evolutionary units with a differentiated northern and southern distribution that correlates basically with the biogeographical regions occupied by this fern (Eurosiberian, Mediterranean). The main haplotypes from both units coexist in the Cantabrian Cornice, and the southern H-II extends, infrequently, northwards along the European Atlantic coast. A similar pattern was discerned by Rumsey *et al*. (1996) for the distribution of the only one polymorphic position detected by PCR-RFLP in the plastid *trn*L-F region.

Our results suggest that this structuring was due to the action of pre-Pleistocene historic rather than biogeographical factors. These are consistent with: 1) long persistence, during Neogene, in independent refugia of the Eurosiberian Atlantic and Mediterranean coasts (considering Madeira and the Canary Islands as part of the Mediterranean Region; Rivas-Martínez *et al*., 2004) and independent evolution of gene pools, and 2) North-South dispersals tracking the Quaternary glacial cycles.

The initial differentiation of *V. speciosa* gene lineages during the Miocene-Pliocene transition (Fig. 3) suggests that the strong aridification at the end of the Miocene, which culminated with the Messinian Salinity Crisis (MSC: 5.96-5.33 Ma; Krijgsman, 1999), caused the contraction of the area, population fragmentation and isolation, and the North-South genetic differentiation of *V. speciosa*. The laurophyllous vegetation (with which *V. speciosa* is associated) was forced to retreat towards moist locations of the Mediterranean and Black sea basins, and the Macaronesian archipelagos (Axelrod, 1975; Kovar-Eder *et al*., 2006; Utescher *et al*., 2007). Moreover, the Atlantic influence on the western end of Europe probably buffered the sharp climate change, allowing the survival of *V. speciosa* and other relict Tertiary plants (Axelrod, 1975; Kovar-Eder *et al*., 2006; Postigo Mijarra *et al*., 2009). A similar pattern has been suggested for *Davallia canariensis* (Liu & Schneider, 2013). The establishment of the Mediterranean climate (3.2 Ma; Suc, 1984) probably favoured the increase in the differentiation between the Eurosiberian and Mediterranean populations, due to the imposed climate constraints for range expansion, especially on the Mediterranean populations. According to Benito Garzón & Sainz Ollero (2002), *V. speciosa* is one of the Tertiary ferns that survived in Iberian and Macaronesian shelters to the change

towards the Mediterranean climate. It is plausible that the Mediterranean climate greatly reduced the Mediterranean populations, generating strong isolation and lowered genetic diversity by genetic drift.

Unlike the Flora Lusitanica (see Beatty & Provan, 2013, 2014), our results do not support the idea that the northern populations originated from those of the south by post-glacial recolonization. The distribution model of genetic diversity shows the concentration of diversity along the Atlantic coast, with intermediate levels in most southern regions and higher levels in most northern regions (Table 1; Fig. 2a). Moreover, the presence of private haplotypes in regions of the northern unit, such as the Azores and Ireland, suggests the long presence of *V. speciosa* in these areas. These private haplotypes are derived from the H-III haplotype (Figs. 2b, 4c), whose diversification started about Pliocene-Pleistocene transition, supporting the presence of *V. speciosa* in these regions before the glaciations (Fig. 3) and its persistence during the Quaternary.

4.4. The Cantabrian Cornice could act as a Tertiary refuge and as a suture zone during the Quaternary

Our findings suggest the Cantabrian Cornice as a main refuge for *V. speciosa* during the Pliocene, in addition to the Pleistocene refuge traditionally accepted for ferns and other plants (Jermy, 1984; Ramil-Rego & Gómez-Orellana, 1996; Vogel *et al*., 1999a; Hunt *et al*., 2009; Beatty & Provan, 2013, 2014). This area harbours an endemic haplotype and the highest haplotype diversity (for cpDNA; Fig. 2a,b; Table 1), evidencing long-term occupation (Trewick *et al*., 2002) and demographic stability (Fig. 4a).

On the other hand, the mixture of haplotypes from the south and the north on the Cantabrian coast also suggests a complementary hypothesis—that is, this region might have been a more recent secondary contact area between the two lineages. Our results imply that this contact happened by the southward expansion of the northern lineage (evidenced by BSP and MASCOT analyses; Fig. 4a-c), in response to changes in habitat suitability caused by glacial cycles (SDM suggests a migration route during the last glacial period; Fig. 5c). The southern lineage could have survived the glacial period along the entire Atlantic coastal strip of the Iberian Peninsula (Fig. 5c) and may have separated from the Cantabrian Cornice during the post-glacial period (Fig. 5b), when

the Mediterranean climate returned along the entire western and southern Iberian coast. This would have left the species in small, favourable climatic pockets, as occurs at present, under the effect of strong genetic drift (Ben-Menni Schuler *et al*., 2017). Rumsey *et al*. (2005) has also suggested this coastal connection and its fragmentation by a process of aridification.

Finally, the MASCOT analysis also suggests migrations of H-I and H-II into northern regions (Figs. 4b,c). Given the demographic stability of the southern and Cantabrian regions (Fig. 3a; Table S3), these results might reflect long-distance dispersal events.

4.5. Central and North Europe

One of the most intriguing biogeographical issues of *V. speciosa* concerns when, how, and from where the gametophyte reached its wide continental and Northern European distribution. Our SDM results show a strong effect of the glaciations on those populations (Fig. 5c), supporting the hypothesis of post-glacial colonization from southern and/or western refugia (proposed by Rumsey *et al*., 1998b, 1999), against the hypothesis of spreading from more northerly Tertiary and peri-glacial refugia (not ruled out by the aforementioned authors). Moreover, the almost total absence of cpDNA diversity in the Central European populations agrees with the post-glacial colonization (Table 1). However, our data do not determine exactly when these populations might have become established and thus do not support or refute the post-glacial hypsithermal period (c. 8.000-4.500 yrs. B.P.) suggested by Rumsey *et al*. (1998b, 1999).

The haplotypes we found strongly differentiate Czech Republic populations from those of Luxembourg and the Vosges (Fig. 2a), suggesting at least two independent long-distance dispersal events. In the Czech Republic, we found only H-VI, suggesting Ireland and/or the Cantabrian Cornice as possible source areas. The presence of H-I in the Vosges connect this region with the southern refugia (Andalusia) or with the Cantabrian Cornice. Molecular markers with more resolution could help to increase the accuracy of these analyses (Ben-Menni Schuler *et al*., In prep.). The *gapCp* sequences were not useful since they did not show any structure.

4.6. Macaronesian archipelagos

Our results indicate the arrival of *V. speciosa* to Macaronesia through different dispersal events to different areas and at different times. The haplotype distribution shows the phylogeographical relationships between the Canary/Madeira archipelagos and southern Iberia, and the relationship between the Azores and north-western Europe, suggesting different sources of colonization for these archipelagos. A similar pattern to has been observed in *Asplenium hemionitis* (Durães, 2005). The floristic link between the Canary Islands and Madeira, and the southern Iberia have been proposed by several authors (e.g. Sunding, 1979; de Nicolás *et al*., 1989; Vanderpoorten *et al*., 2007), while other researchers have differentiated the flora of the Azores from that of the Canary Islands and Madeira, ascribing them to different biogeographical regions (e.g. Lobin, 1982, Rivas-Martínez *et al*., 2004). Geological evidence indicates possible ancient land connections between the Canary Islands, Madeira, and southern Iberia (by large volcanic seamounts; García-Talavera, 1999; Fernández-Palacios *et al*. 2011 and references therein), which are thought to have contributed substantially to the configuration of Macaronesian biota (e.g. van den Broeck *et al*., 2008), perhaps accounting for the haplotype pattern observed for *V. speciosa* in the Macaronesia.

The colonization of Macaronesia could have occurred during the Plio-Pleistocene, as has been confirmed for relict angiosperms (Kondraskov *et al*., 2015). *Vandenboschia speciosa* reached the Canary Islands at least during the interglacial Donau-Günz, when the H-VIII haplotype, private to the Canary Islands, diverged from H-II (1.3 Ma; Fig. 3), and probably from Madeira (with laurel forest since at least 1.8 Ma; Góis-Marques *et al*., 2018a) or from the south of the Iberian Peninsula.

On the other hand, *V. speciosa* could have reached the Azores at the end of the Pliocene, before the glacial periods. Our data support the idea of a process of independent diversification of the four haplotypes derived from H-III (Figs. 2b, 3, Fig. S3), including the two private of Azores, and that started 2.58 Ma (Fig. 3; external nodes in the H-III haplogroup lacks of bootstrap support). This would place H-III in the Azores before that date, although a more recent origin of these private haplotypes cannot be ruled out. Góis-Marques *et al*. (2018b) concluded a Pleistocene to Holocene age for all plant fossils described in the Azores. *Vandenboschia speciosa* could have

arrived to the Azores at Santa Maria Island (~6 Ma), the only island older than 1.5 Ma (reviewed by Góis-Marques *et al*., 2018b). From there, *V. speciosa* should have spread to the other islands. The haplotype sharing and the signs of recent population expansion (Table S3) supports the contention of the intra-archipelago dispersal of *V. speciosa* in the Azores. The presence in the Azores of haplotype H-VI suggests the immigration of the species from Cantabrian Cornice and/or from Ireland.

4.7. Implications for conservation

Vandenboschia speciosa, a species considered threatened at the European level (Anonymous, 1979, 1992), is listed in the national red lists of many countries (with different status, from Not Threatened to Critically Endangered; Christenhusz *et al*., 2017). However, no large-scale common conservation plan covers the species throughout its distribution area. Our contribution constitutes a baseline study for a possible conservation plan for the species at the European level. The identification of two evolutionary units (north/south), of the absence of genetic structuring between the phases of the life cycle (gametophyte can be regarded as a 'seed-bank' and genetic reservoir for the species; Rumsey *et al*., 1999), and the quantification of genetic diversity throughout the area of the species could improve the conservation and management of this fern.

In relation to the future, the projection of the SDM under 2080 climate predictions suggests a northward migration of *V. speciosa* and a high risk that the southern evolutionary unit could disappear (Fig. 5d, Fig. S5). The only areas that are maintained with steady potential are the UK and Ireland. This is consistent with the forecasts made for *Asplenium fontanum* and other ferns (Robinson, 2009; Bystriakova *et al*., 2013). The expected reduction of the genetic diversity makes it necessary to intensify the conservation measures of the populations farther south.

Conclusions

Our results show a Tertiary origin for *V. speciosa* and the presence of two principal evolutionary units with a differentiated North (Eurosiberian Region) and South (Mediterranean Region) distribution. The strong aridification at the end of the Miocene

that culminated with the Messinian Salinity Crisis is suggested to be the cause of this differentiation. The Cantabrian Cornice must have been a main refuge for *V. speciosa* during the Tertiary and the Pleistocene glaciations, but it could also have been a recent contact zone between the two lineages due to expansions of the area from the north towards the south and from the Cantabrian towards the north following the glacial-interglacial cycles. Current central European populations appear to have derived from post-glacial multiple long-distance dispersals. Our results indicate the arrival of *V. speciosa* to Macaronesia through different dispersal events and reveal the phylogeographical relationships between the Canary/Madeira archipelagos and southern Iberia, as well as between the Azores and the north-western Europe.

References

Anonymous. (1979). Convention of the conservation of European wildlife and natural habitats. *European Treaty Series*, **104**, 1–27.

Anonymous. (1992). Council Directive 92/43/EEC of 21 May (1992) on the conservation of habitats and wild fauna and flora. *Official Journal of the European Communities L*, **206**, 7–50.

Avise, J. C. (2000). *Phylogeography. The history and formation of species*. Cambridge, MA: Harvard University Press.

Axelrod, D. I. (1975). Evolution and biogeography of Madrean-Tethyan sclerophyll vegetation. *Annals of the Missouri Botanical Garden*, **62**, 280–334.

Barrón, E. (2003). Evolución de las floras terciarias en la Península Ibérica. *Monografías del Jardín Botánico de Córdoba*, **11**, 63–74.

Barrón, E., & Peyrot, D. (2006). La vegetación forestal en el Terciario. In J. Carrión, S. Fernández & N. Fuentes (Eds.), *Paleoambientes y cambio climático* (pp. 56–77). Murcia, Spain: Fundación Séneca/Agencia de Ciencia y Tecnología de la Región de Murcia.

Beatty, G. E., & Provan, J. (2013). Post-glacial dispersal, rather than in situ glacial survival, best explains the disjunct distribution of the Lusitanian plant species *Daboecia cantabrica* (Ericaceae). *Journal of Biogeography*, **40**, 335–344.

Beatty, G. E., & Provan, J. (2014). Phylogeographical analysis of two cold-tolerant plants with disjunct Lusitanian distributions does not support in situ survival during the last glaciation. *Journal of Biogeography*, **41**, 2185–2193.

Ben-Menni Schuler, S., García-López, M. D. C., Lopez-Flores, I., Nieto-Lugilde, M., & Suárez-Santiago, V. N. (2017). Genetic diversity and population history of the Killarney fern, *Vandenboschia speciosa* (Hymenophyllaceae), at its southern distribution limit in continental Europe. *Botanical Journal of the Linnean Society*, **183**, 94–105.

Benito Garzón, M., & Sainz de Ollero, H. (2002). *Potencialidad del elemento paleotropical lauroide en la Península Ibérica.* MSc thesis. Madrid, Spain: Universidad Autónoma de Madrid.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., … Drummond, A. J. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, **10,** e1003537.

Bozukov, V. (2008). Genus *Trichomanes* in the European and Bulgarian Cenozoic macroflora. *Phytologia Balcanica*, **14**, 3–9.

van den Broeck, H., Breugelmans, K., de Wolf, H., & Backeljau, T. (2008). Completely disjunct mitochondrial DNA haplotype distribution without a phylogeographic break in a planktonic developing gastropod. *Marine Biology*, **153**, 421–429.

Bystriakova, N., Ansell, S. W., Russell, S. J., Grundmann, M., Vogel, J. C., & Schneider, H. (2014). Present, past and future of the European rock fern *Asplenium fontanum*: combining distribution modelling and population genetics to study the effect of climate change on geographic range and genetic diversity. *Annals of Botany*, **113**, 453–465.

Calleja, J. A., Benito Garzón, M., & Sáinz Ollero, H. (2009). A Quaternary perspective on the conservation prospects of the Tertiary relict tree *Prunus lusitanica* L. *Journal of Biogeography*, **36**, 487–498.

Chen, C., Qi, Z. C., Xu, X. H., Comes, H. P., Koch, M.A., Jin, X. J., … Qiu, Y. X. (2014). Understanding the formation of Mediterranean–African–Asian disjunctions: evidence for Miocene climate-driven vicariance and recent long-

distance dispersal in the Tertiary relict *Smilax aspera* (Smilacaceae). *New Phytologist*, **204**, 243–255.

Christenhusz, M., Bento Elias, R., Dyer, R., Ivanenko, Y., Rouhan, G., Rumsey, F., & Väre, H. (2017). *Trichomanes speciosum*. *The IUCN Red List of Threatened Species*, **2017**, e.T162169A85429885.

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Cronk, Q. C.B. (1992). Relict floras of Atlantic islands: patterns assessed. *Biological Journal of the Linnean Society*, **46**, 91–103.

De Nicolas, J. P., Fernández-Palacios, J. M., Ferrer, F. J., & Nieto, E. (1989). Inter-island floristic similarities in the Macaronesian region. *Vegetatio*, **84**, 117–125.

Drummond, A. J., Ho, S. Y., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, e88.

Drummond, A. J., Rambaut, A., Shapiro, B. E. T. H., & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, **22**, 1185–1192.

Durães, C. J. (2005). Genetic variation and evolutionary history of the Macaronesian fern *Asplenium hemionitis* L. PhD thesis, Edinburgh University, Edinburgh, UK.

Ebihara, A., Ishikawa, H., Matsumoto, S., Lin, S. J., Iwatsuki, K., Takamiya, M., … Ito, M. (2005). Nuclear DNA, chloroplast DNA, and ploidy analysis clarified biological complexity of the *Vandenboschia radicans* complex (Hymenophyllaceae) in Japan and adjacent areas. *American Journal of Botany*, **92**, 1535–1547.

Ebihara, A., Matsumoto, S., & Ito, M. (2009). Hybridization involving independent gametophytes in the *Vandenboschia radican*s complex (Hymenophyllaceae): a new perspective on the distribution of fern hybrids. *Molecular Ecology*, **18**, 4904–4911.

Eichler, M., & Kempf, M. (2012). *Nachuntersuchungen zur Verbreitung des Prächtigen Dünnfarns (Trichomanes speciosum) in Hessen (Art des Anhangs II der FFH-*

Richtlinie) im Jahr (2011) – Endbericht (2012) –. Hessen, Germany: Hessen-Forst FENA Naturschutz.

Erdei, B., Hably, L., Kázmér, M., Utescher, T., & Bruch, A. A. (2007). Neogene flora and vegetation development of the Pannonian domain in relation to palaeoclimate and palaeogeography. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **253**, 115–140.

Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Fernández-Palacios, J. M, de Nascimento, L., Otto, R., Delgado, J. D., García-del-Rey, E., Arévalo, J. R., & Whittaker, R. J. (2011). A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography*, **38**, 226–246.

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Gascuel, O. (1997). BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Molecular Biology and Evolution*, **14**, 685–695.

García-Talavera, F. (1999). La Macaronesia. Consideraciones geológicas, biogeográficas y palaeoecológicas. In J. M. Fernández-Palacios, J. J. Bacallado & J. A. Belmonte (Eds.), *Ecología y cultura en Canarias* (pp. 39–63). Santa Cruz de Tenerife, Canary Islands, Spain: Museo de la Ciencia, Cabildo Insular de Tenerife.

García-Verdugo, C., Calleja, J. A., Vargas, P., Silva, L., Moreira, O., & Pulido, F. (2013). Polyploidy and microsatellite variation in the relict tree *Prunus lusitanica* L.: how effective are refugia in preserving genotypic diversity of clonal taxa? *Molecular Ecology*, **22**, 1546–1557.

Góis-Marques, C. A., Madeira, J., & Menezes de Sequeira, M. (2018a). Inventory and review of the Mio-Pleistocene São Jorge flora (Madeira Island, Portugal): palaeoecological and biogeographical implications. *Journal of Systematic Palaeontology,* **16**, 159–177.

Góis-Marques, C. A., de Nascimento, L., Menezes de Sequeira, M., Fernández-Palacios, J. M., & Madeira, J. (2018). The Quaternary plant fossil record from the volcanic Azores Archipelago (Portugal, North Atlantic Ocean): a review. *Historical Biology*, 1–17.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology,* **59**, 307–321.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series,* **41**, 95–98.

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society,* **58**, 247–276.

Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society,* **68**, 87–112.

Hewitt, G. M. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907.

Hewitt, G. M. (2003). Ice ages: their impact on species distri-butions and evolution. In L. J. Rothschild, A. M. Lister (Eds.), *Evolution on planet Earth* (pp. 339–361). New York, USA: Academic Press.

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: A Journal of the Royal Meteorological Society,* **25**, 1965–1978.

Hunt, H. V., Ansell, S. W., Russell, S. J., Schneider, H., & Vogel, J. C. (2009). Genetic diversity and phylogeography in two diploid ferns, *Asplenium fontanum* subsp.

fontanum and *A. petrarchae* subsp. *bivalens*, in the western Mediterranean. *Molecular Ecology*, **18**, 4940–4954

Ivanov, D., Utescher, T., Mosbrugger, V., Syabryaj, S., Djordjević-Milutinović, D., & Molchanoff, S. (2011). Miocene vegetation and climate dynamics in eastern and central Paratethys (southeastern Europe). *Palaeogeography, Palaeoclimatology, Palaeoecology*, **304**, 262–275.

Jermy, A. C. (1984). Origin and distribution of Pteridophytes in the Mediterranean area. *Webbia*, **38**, 397–416.

Jiménez, A., Holderegger, R., Csencsics, D., & Quintanilla, L. G. (2010). Microsatellites reveal substantial among-population genetic differentiation and strong inbreeding in the relict fern *Dryopteris aemula*. *Annals of Botany*, **106**, 149–155.

Kondraskov, P., Schütz, N., Schüßler, C., de Sequeira, M. M., Guerra, A. S., Caujapé-Castells, J., … (2015). Biogeography of Mediterranean Hotspot Biodiversity: Re-Evaluating the 'Tertiary Relict' Hypothesis of Macaronesian Laurel Forests. *PLoS One*, **10**, e0132091.

Kovar-Eder, J., Kvaček, Z., Martinetto, E., & Roiron, P. ((2006)). Late Miocene to Early Pliocene vegetation of southern Europe (7–4 Ma) as reflected in the megafossil plant record. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **238**, 321–339.

Krijgsman, W., Hilgen, F. J., Raffi, I., Sierro, F. J., & Wilson, D. S. (1999). Chronology, causes and progression of the Messinian salinity crisis. *Nature*, **400**, 652.

Krukowski, M., & Świerkosz, K. (2004). Discovery of the gametophytes of *Trichomanes speciosum* (Hymenophyllaceae: Pteridophyta) in Poland and its biogeographical importance. *Fern Gazette*, **17**, 79–84.

Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Liu, H., & Schneider, H. (2013). Evidence supporting *Davallia canariensis* as a Late Miocene relict endemic to Macaronesia and Atlantic Europe. *Australian Systematic Botany*, **26**, 378–385.

Lobin, W. (1982). Untersuchung uber Flora, Vegetation und biogeographische Beziechungen der Kapverdischen Inseln. *Courier Forschungsinstitut Senckenberg*, **53**, 1–112.

Loriot, S., & Geslin, J. (2005). *Trichomanes speciosum* Willd. (Hymenophylacceae, Pteridophyta) dans le Massif Armoricain. E.R.I.C.A., **19**, 3–22.

Louis-Arsène, F. (1953). *Trichomanes speciosum* Willd. en Bretagne. *Bulletin de la Société Botanique de France,* **100**, 6.

Maccagni, A., Parisod, C., & Grant, J. R. (2017). Phylogeography of the moonwort fern *Botrychium lunaria* (Ophioglossaceae) based on chloroplast DNA in the Central-European Mountain System. *Alpine Botany,* **127**, 185–196.

Mai, D. H. (1989). Development and regional differentiation of the European vegetation during the Tertiary. *Plant Systematics and Evolution,* **162**, 79–91.

Mai, D. H. (1991). Palaeofloristics changes in Europe and the confirmation of the Arctotertiary–Palaeotropical geofloral concept. *Review of Palaoeobotany and Palynology*, **68**, 29–36.

Makgomol, K., & Sheffield, E. (2001). Gametophyte morphology and ultrastructure of the extremely deep shade fern, *Trichomanes speciosum*. *New Phytologist*, **151**, 243–255.

Makgomol, K., & Sheffield, E. (2005). Development of gametophytes from gemmae of Killarney Fern (*Trichomanes speciosum* Willd., Hymenophyllaceae, Pteridophyta). *Fern Gazette,* **17**, 163.

Manton, I. (1950). Problems of cytology and evolution in the Pteridophyta. Cambridge, UK: Cambridge University Press.

Marchetti, D. (2002). Notule pteridologiche italiche. I (1-31). *Annali del Museo Civico di Rovereto, Sezione: Archeologia, Storia, Scienze Naturali,* **16**, 371–392.

Martínez, B., Viejo, R. M., Carreño, F., & Aranda, S. C. (2012). Habitat distribution models for intertidal seaweeds: responses to climatic and non-climatic drivers. *Journal of Biogeography*, **39**, 1877–1890.

Miao, Y., Herrmann, M., Wu, F., Yan, X., & Yang, S. (2012). What controlled Mid–Late Miocene long-term aridification in Central Asia? — Global cooling or Tibetan Plateau uplift: A review. *Earth-Science Reviews*, **112**, 155–172.

Molnar, P., Boos, W. R., & Battisti, D. S. (2010). Orographic controls on climate and paleoclimate of Asia: thermal and mechanical roles for the Tibetan Plateau. *Annual Review of Earth and Planetary Sciences*, **38**, 77–102.

Müller, N. F., Rasmussen, D., & Stadler, T. (2018). MASCOT: Parameter and state inference under the marginal structured coalescent approximation. *Bioinformatics*, **34**, 3843-3848.

Ní Dhúill, E. (2014). Conservation Biology of the Threatened Killarney Fern (*Trichomanes speciosum* Willd.) in Ireland. PhD thesis. Dublin, Ireland: University of Dublin, Trinity College.

Nitta, J. H., Ebihara, A., & Ito, M. (2011). Reticulate evolution in the *Crepidomanes minutum* species complex (Hymenophyllaceae). *American Journal of Botany*, **98**, 1782–1800.

Palamarev, E., & Petkova, A. (1990). The Paleogene macroflora of the Rhodopes region. I. Polypodiophyta–Polypodiopsida. *Fitologiya*, **38**, 3–22.

Petit, R. J, Hampe, A., & Cheddadi, R. (2005). Climate change and tree phylogeography in the Mediterranean. *Taxon*, **54**, 877–885.

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

Pichi-Sermolli, R. E. G. (1979). A survey of the pteridological flora of the Mediterranean region. *Webbia*, **34**, 175–242.

Pichi Sermolli, R. E. G. (1991). On the taxonomy and nomenclature of some species from tropical Africa of the genus *Triplophyllum* Holttum (Dryopteridaceae). *Webbia*, **45**, 117–135.

Pinson, J. B., Chambers, S. M., Nitta, J. H., Kuo, L. Y., & Sessa, E. B. (2017). The separation of generations: biology and biogeography of long-lived sporophyteless fern gametophytes. *International Journal of Plant Sciences*, **178**, 1–18.

Postigo Mijarra, J. M., Barrón, E., Gómez Manzaneque, F., & Morla, C. (2009). Floristic changes in the Iberian Peninsula and Balearic Islands. *Journal of Biogeography*, **36**, 2025–2043.

Provan, J., & Bennett, K. D. (2008). Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution*, **23**, 564–571.

Quintanilla, L. G., Pajarón, S., Pangua, E., & Amigo, J. (2007). Allozyme variation in the sympatric ferns *Culcita macrocarpa* and *Woodwardia radicans* at the northern extreme of their ranges. *Plant Systematics and Evolution*, **263**, 135–144.

Ramil-Rego, P., & Gómez-Orellana, L. (1996). Dinámica climática y biogeográfica del área litoral-sublitoral de Galicia. In P. Ramil Rego, C. Fernández Rodríguez & M. Rodríguez Guitián (Eds.), *Biogeografía Pleistocena-Holocena de la Península Ibérica* (pp. 43–71). Santiago de Compostela, Spain: Xunta de Galicia.

Rivas-Martínez, S., Penas, A., & Díaz, T. E. (2004). *Bioclimatic and biogeographic maps of Europe*. [WWW document] URL http://www.globalbioclimatics.org/form/maps.htm

Robinson, R. C. (2009). Invasive and problem ferns: a European perspective. *International Urban Ecology*, **4**, 83–90.

Rodríguez-Sánchez, F., & Arroyo, J. (2008). Reconstructing the demise of Tethyan plants: climate-driven range dynamics of *Laurus* since the Pliocene. *Global Ecology and Biogeography*, **17**, 685–95.

Rodríguez-Sánchez, F., Guzmán, B., Valido, A., Vargas, P., & Arroyo, J. (2009). Late Neogene history of the laurel tree (*Laurus* L., Lauraceae) based on phylogeographical analyses of Mediterranean and Macaronesian populations. *Journal of Biogeography*, **36**, 1270–1281.

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics,* **19**, 1572–1574.

Rumsey, F. J., Barrett, J. A., Gibby, M., Russell, S. J., & Vogel, J. C. (2005). Reproductive strategies and population structure in the endangered pteridophyte *Trichomanes speciosum* (Hymenophyllaceae: Pteridophyta). *Fern Gazette,* **17**, 205.

Rumsey, F. J., Jermy, A. C., & Sheffield, E. (1998a). The independent gametophytic stage of *Trichomanes speciosum* Willd. (Hymenophyllaceae), the Killarney Fern and its distribution in the British Isles. *Watsonia,* **22**, 1–20.

Rumsey, F. J., & Sheffield, E. (1996). Inter-generational ecological niche separationand the "independent gametophyte" phenomenom. In J. M. Camus, M. Gibby & R. J. Johns (Eds.), *Pteridology in perspective* (pp. 563–570). London, UK: Royal Botanic Gardens, Kew.

Rumsey, F. J., Vogel, J. C., Russell, S. J., Barrett, J. A., & Gibby, M. (1998b). Climate, colonisation and celibacy: population structure in Central European *Trichomanes speciosum* (Pteridophyta). *Botanica Acta,* **111**, 481–489.

Rumsey, F. J., Vogel, J. C., Russell, S. J., Barrett, J. A., & Gibby, M. (1999). Population structure and conservation biology of the endangered fern *Trichomanes speciosum* Willd. (Hymenophyllaceae) at its northern distributional limit. *Biological Journal of the Linnean Society,* **66**, 333–344.

Sánchez-Velázquez, T. (2013). *Trichomanes speciosum* (Pteridophyta Hymenophyllaceae) en la isla de Gran Canaria, Islas Canarias. *Botánica Macaronésica,* **28**, 82–92.

Sanz-Azcue, I., Díez-López, J., & Olariaga-Ibarguren, I. (2013). Inventory and mapping of red-listed vascular flora in Hernani municipaly (Gipuzkoa, Basque Country). *Munibe Ciencias Naturales,* **61**, 7–32.

Schuettpelz, E., & Pryer, K. M. (2009). Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. *Proceedings of the National Academy of Sciences,* **106**, 11200–11205.

Stewart, J. R., & Lister, A. M. (2001). Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology & Evolution*, **16**, 608–613.

Suc, J. P. (1984). Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature*, **307**, 429.

Sunding, P. (1979)**.** Origins of the Macaronesian fauna. In D. Bramwell (Ed.), *Plants and Islands* (pp. 13–40). London, UK: Academic Press.

Svenning, J. C. (2003). Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. *Ecology Letters*, **6**, 646–653.

Svenning, J. C, & Skov, F. (2007). Could the tree diversity pattern in Europe be generated by postglacial dispersal limitation? *Ecology Letters*, **10**, 453–460.

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cossons, J. F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453– 464.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Templeton, A. R., Crandall, K. A., & Sing, C. F. (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619– 633.

Trewick, S. A., Morgan-Richards, M., Russell, S. J., Henderson, S., Rumsey, F. J., Pintér, I., … Vogel, J. C. (2002). Polyploidy, phylogeography and Pleistocene refugia of the rockfern *Asplenium ceterach*: evidence from chloroplast DNA. *Molecular Ecology*, **11**, 2003–2012.

Utescher, T., Erdei, B., François, L., & Mosbrugger, V. (2007). Tree diversity in the Miocene forests of Western Eurasia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **253**, 226–250.

Van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., … (2011). The representative concentration pathways: an overview. *Climatic Change*, **109**, 5.

Vanderpoorten, A., Rumsey, FJ., & Carine, M. A. (2007). Does Macaronesia exist? Conflicting signal in the bryophyte and pteridophyte floras. *American Journal of Botany,* **94**, 625–639.

Vogel, J. C., Rumsey, F. J., Russell, S. J., Cox, C. J., Holmes, J. S., Bujnoch, W., … Gibby, M. (1999b). Genetic structure, reproductive biology and ecology of isolated populations of *Asplenium csikii* (Aspleniaceae, Pteridophyta). *Heredity*, **83**, 604–612.

Vogel, J. C., Rumsey, F. J., Schneller, J. J., Barrett, J. A., & Gibby, M. (1999a)**.** Where are the glacial refugia in Europe? Evidence from pteridophytes. *Biological Journal of the Linnean Society,* **66**, 23–37.

Supporting Information

Fig. S1 Maximum likelihood tree for the *gapCp* sequences of *Vandenboschia speciosa* determined with PhyML; numbers above branches are support values from the Shimodaira-Hasegawa-Like implementation of the approximate likelihood-ratio test; the colours differentiate the two identified paralogue copies; on the right are the names of the paralogues, the number of sequences of each one found by us in *V. birmanica* and *V. davallioides*, and the names of the sequences of Ebihara *et al*. (2005).

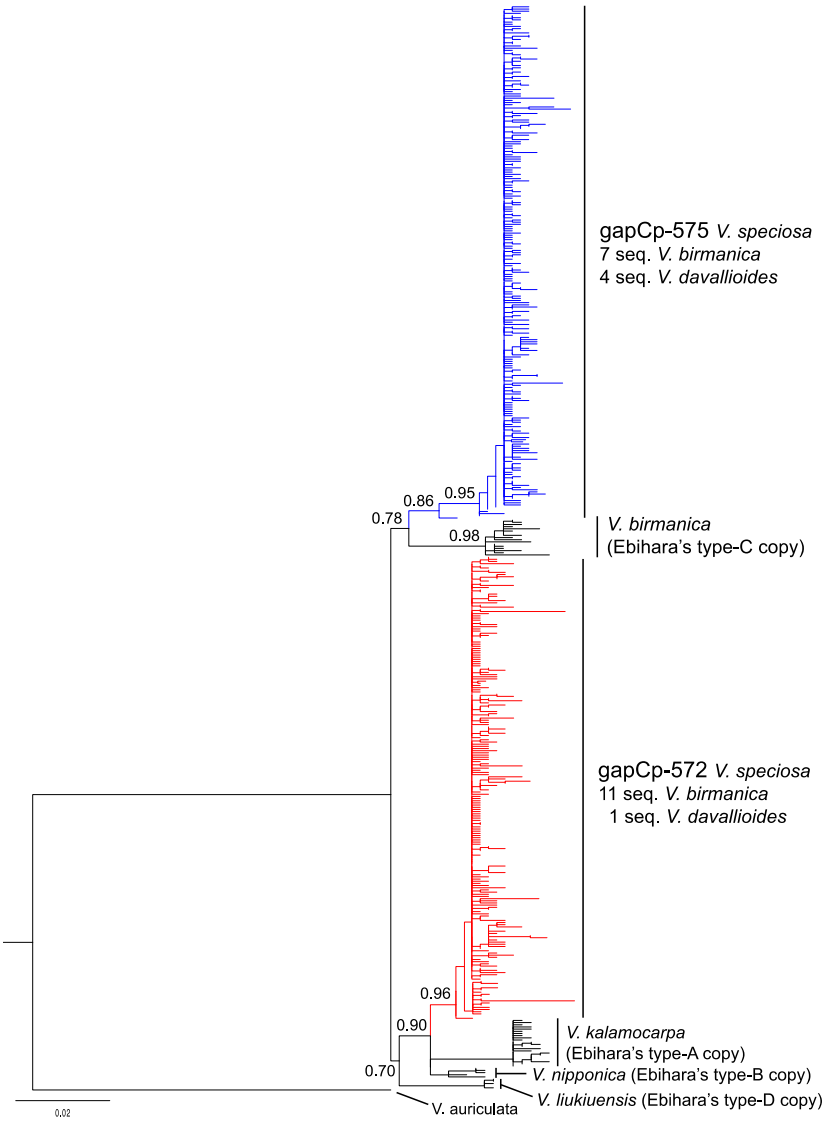


Fig. S2 Statistical parsimony networks of the *gapCp* sequences. Networks for the gapCp-572 (upper) and gapCp-575 (lower) copies are shown; the haplotypes are denoted with Roman numerals; numbers above branches represent the positions of the nucleotide substitution in the aligned sequences; circle sizes are proportional to the number of sequences found for each haplotype.

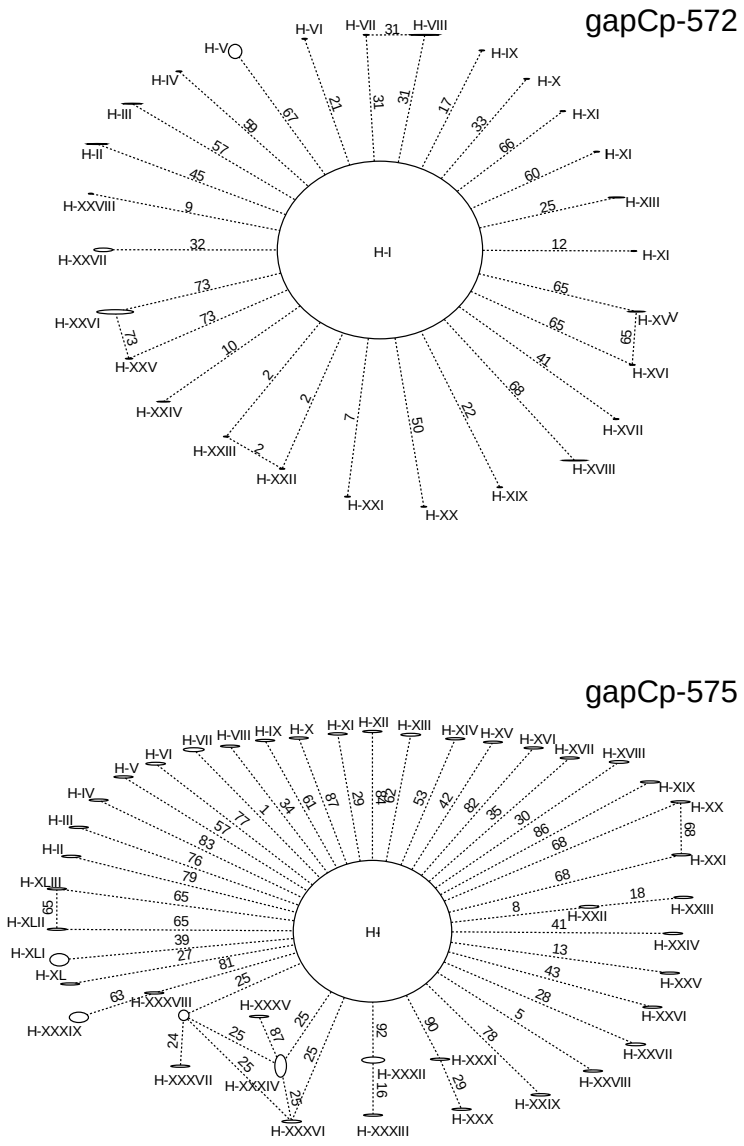


Fig. S3 cpDNA haplotype phylogeny derived from the Bayesian inference with MrBAYES. The tree is 50% majority-rule consensus tree of the sample of trees after discarding 25% as burn-in; haplotypes are denoted with Roman numerals; numbers above branches are Bayesian posterior probabilities; for outgroup species, accession numbers of sequences taken from GenBank are shown before the species name, while the voucher references of sequences determined by us are shown after the species name.

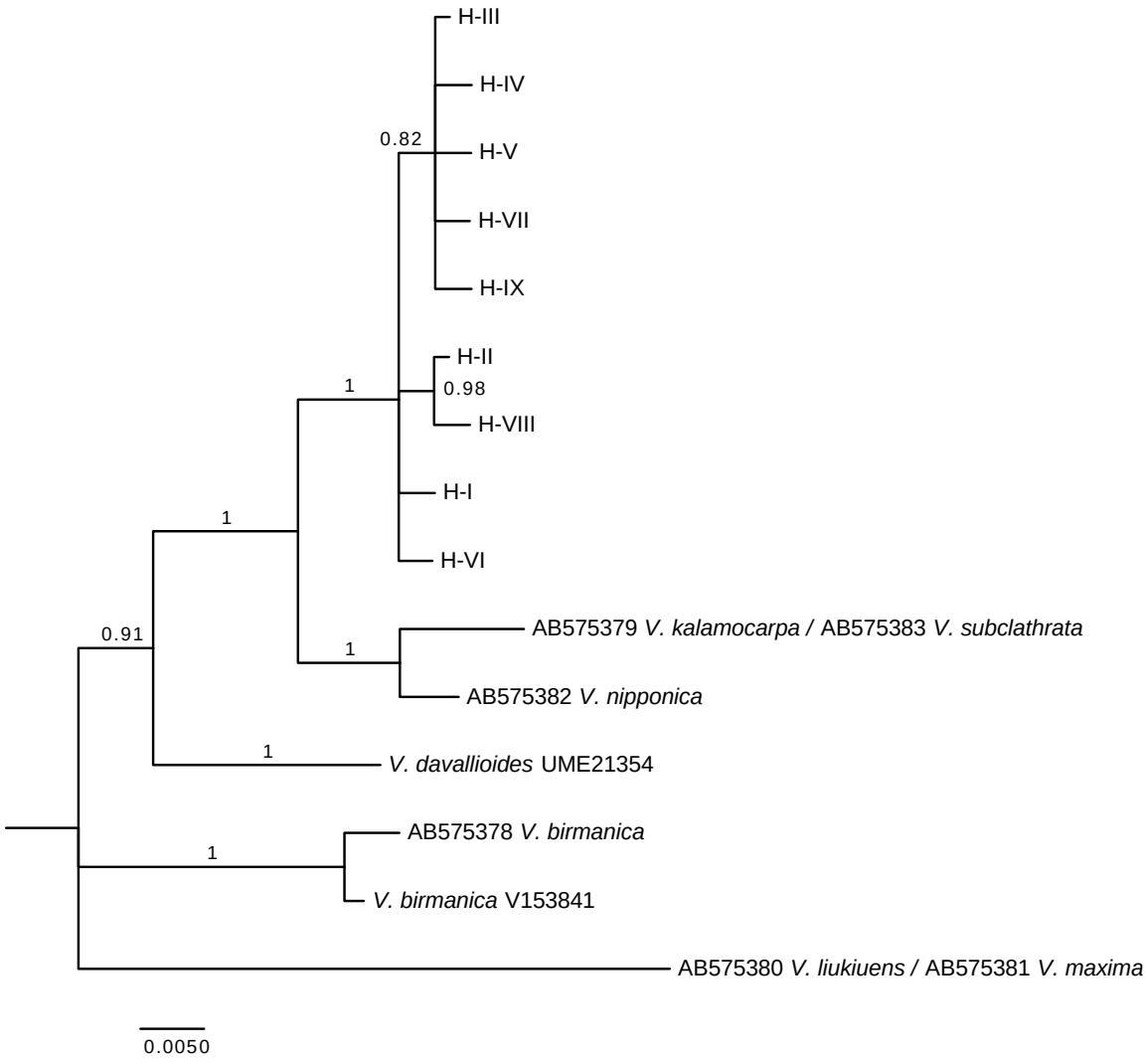


Fig. S4 Diversification-time estimates for the genus *Vandenboschia* and outgroups obtained with BEAST using the *rbc*L dataset. Numbers above branches are the mean divergence ages (in million years ago) for each node; grey bars represent the 95% highest posterior density intervals for each node; the bold uppercase letters A and B indicate the dates used to calibrate the haplotype phylogeny (Fig. 3); numbers after nodes are BEAST posterior probabilities; accession numbers of sequences (all taken from GenBank) are shown before the species name; sequence for *Vandenboschia speciosa* is marked in bold text; the time scale is printed in million years ago (Ma).

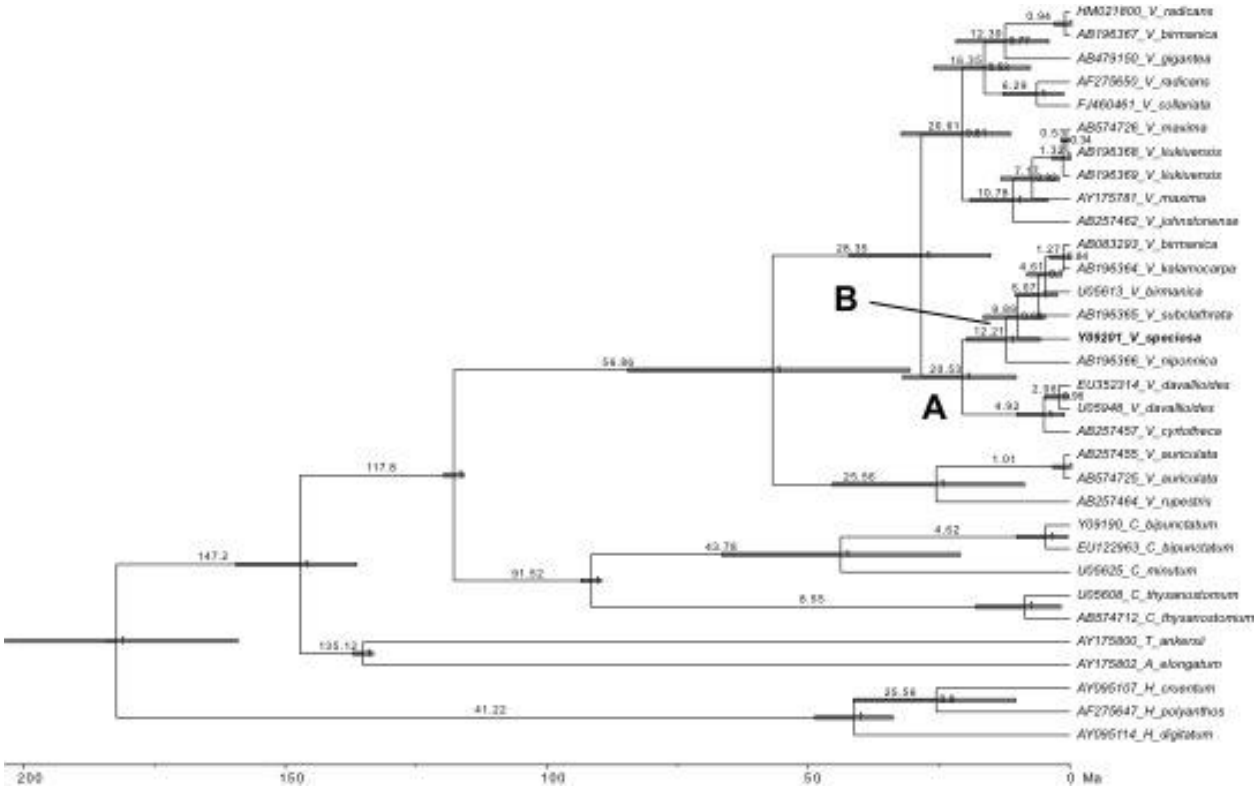


Fig. S5. Results of species distribution modelling for the sporophyte and gametophyte phases of the life cycle of *Vandenboschia speciosa* using the maximum entropy algorithm and the Community Climate System Model (CCSM), as implemented in MAXENT. Projections for current, last interglacial (LIG, *c.* 120-140 kya), last glacial maximum (LGM, *c.* 21 kya), and future (2080) conditions are shown.

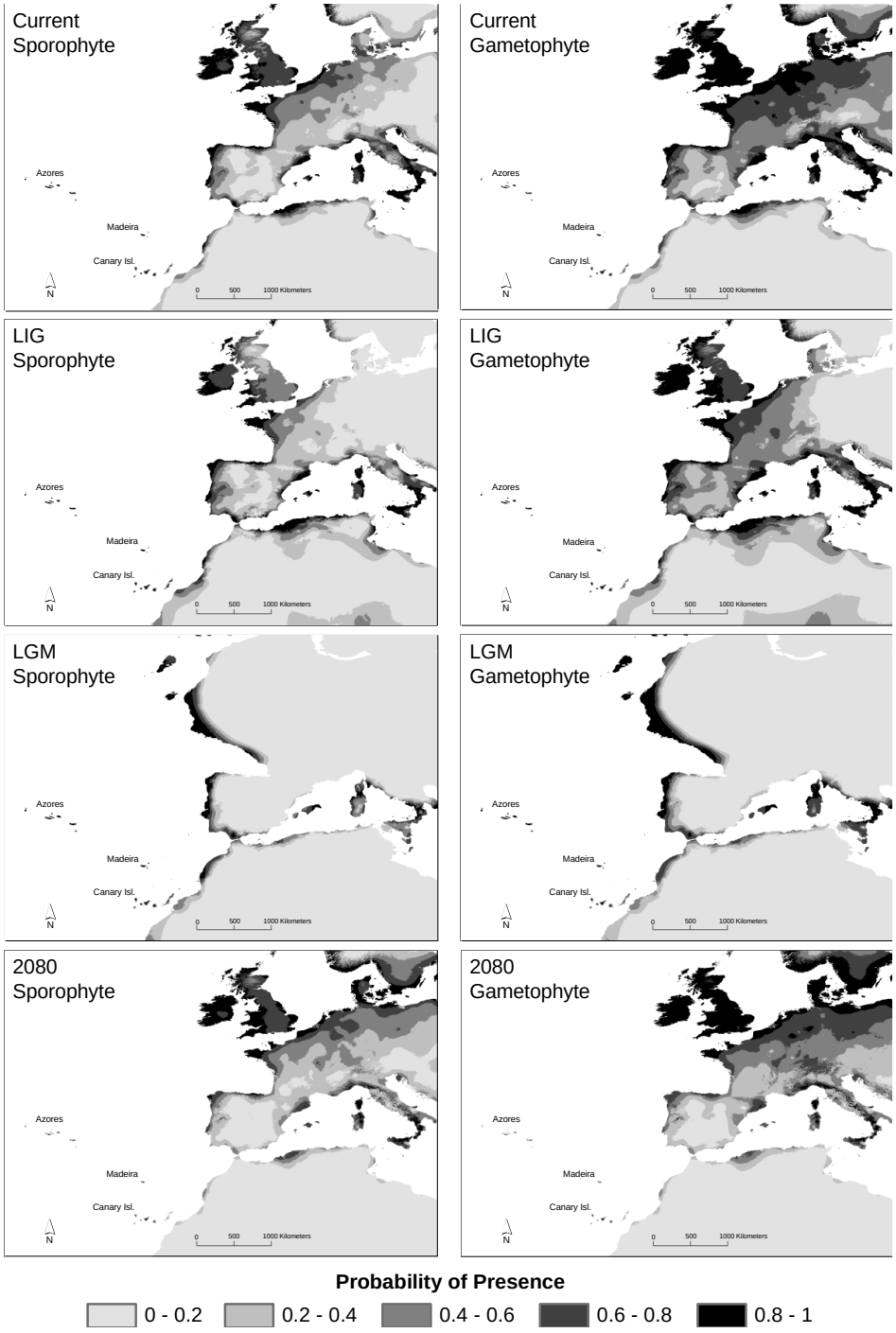


Table S1 Sampling details of *Vandenboschia speciosa* populations and outgroup species used in the present study

Code	Location	Voucher[a]	Geographical coordinates	Sample size *trn*H-*psb*A	Sample size *gapCp* [b]
Andalusia					
ALM	Almoraima		N36.304°/W5.520°	G: 5	5
COQ	Canuto de Ojén Quesada	*GDA 61589*	N36.127°/W5.585°	G:5/S:11	5
CRM	Cabecera del río de la Miel	*GDA 62522*	N36.105°/W5.528°	S:6	
MCH	Moracha	*GDA 62523*	N36.497°/W5.584°	G:5/S:5	5
SCD	Garganta de la Sauceda	*GDA 62524*	N36.535°/W5.605°	S:3	
SDN	Sierra del Niño	*GDA 62525*	N36.186°/W5.610°	S:4	
VIF	Valdeinfierno	*GDA 62526*	N36.224°/W5.604°	G:5/S:5	5
Azores					
CAR	Algar do Carvão		N38.727°/W27.215°	G:5/S:5	5
CID	Sete Cidades		N37.835°/W25.788°	G:5/S:5	5
CON	Lagoa do Congro		N37.754°/W25.407 °	G:5/S:5	5
NAT	Gruta do Natal		N38.738°/W27.264°	G:5/S:5	5
Basque Country					
AZK	Azketa erreka		N43.194/W1.940°	G:5/S:5	4
ERR	Erramundi erreka		N43.377°/W1.826°	G:3/S:5	3
ITU	Iturraingo erreka		N43.373°/W1.833°	G:5	2
USO	Usoko erreka		N43.242°/W1.908°	G:5/S:11	5
Canary Isl.					
ANC	Ancón Negro		N28.134°/W17.273°	G:5/S:5	5
CED	Cedro		N28.120°/W17.225°	G:4/S:5	4
IJU	Ijuana		N28.560°/W16.172°	G:5/S:5	5
PIJ	El Pijaral		N28.553°/W16.188°	G:5/S:5	5
ZAR	Cedro		N28.119°/W17.224°	G:5/S:5	5
Czech Republic					
HAR	Harasov		N50.410°/E14.567°	G:6	3
MUZ	Mužský		N50.528°/E15.054°	G:5	3
SKA	Skalka		N50.585°/E14.424°	G:5	3
Galicia					

EUM	Eume	N43.404°/W8.087°	G:5/S:1	5
SEI	Seixo	N43.706°/W7.946°	G:5/S:5	5
Ire-Wal-Bri*				
COR	Cork	N51.570°/W9.148°	G:5/S:5	5
DEV	Devil's Bridge	N52.376°/W3.849°	G:4	3
LIM	Limerick	N52.663°/W8.387°	G:5/S:4	5
TAU	Taupont	N47.962°/W2.429°	G:2/S:4	2
WAT	Waterford	N52.115°/W7.585°	G:5/S:5	5
Italy				
SER	Seravezza	N44.015°/E10.219°	G:5/S:6	5
STA	Stazzema	N43.991°/E10.315°	G:5	5
Luxembourg				
ARD	Ardennes	N49.906°/E5.954°	G:5	3
BEA	Beaufort	N49.832°/E6.287°	G:3	3
ROL	Rollingen	N49.739°/E6.133°	G:3	3
Madeira				
FRI	Ribeiro Frio	N32.734°/W16.886°	G:5/S:5	5
POR	Levada Portadela	N32.747°/W16.823°	G:3/S:1	3
URZ	Lombo do Urzal	N32.776°/W16.977°	S:5	
Vosges du Nord				
BIT	Bitche	N49.024°/E7.620°	G:5	3
PIE	La Petite Pierre	N48.848°/E7.301°	G:5	3
V. davallioides[c]		*UME 213055*	S:1	1
V. birmanica[c]		*UPS V-153840*; *UPS V-153841*	S:1 S:1	1

[a] Voucher available only for sporophytes; [b] *gapCp* only studied with gametophyte individuals except outgroups; [c] species analysed as outgroups. [*]Ireland-Wales-Brittany. G, Gametophyte individuals; S, sporophyte individuals.

Table S2 Percentage contribution and permutation importance of selected model for the species distribution modelling (SDM).

Variable	Gametophyte		Sporophyte		Species as a whole	
	MaxEnt Percent contribution	MaxEnt Permutation importance	MaxEnt Percent contribution	MaxEnt Permutation importance	MaxEnt Percent contribution	MaxEnt Permutation importance
Min. temperature of coldest month	51.5	75	46.5	82.1	51.5	74.6
Mean diurnal range	40.8	15	41.1	10.9	40.8	13.9
Precipitation of warmest quarter	2.8	0.2	6.4	2.7	2.7	0.3
Precipitation seasonality	2.4	0.6	4.4	2.1	2.3	0.7
max. temperature of warmest month	2.2	5.8	0.6	0	2.3	7.2
Precipitation of coldest quarter	0.2	1.5	0.4	0	0.3	1.3
Precipitation of wettest month	0.1	1.8	0.4	1.8	0.1	2
Precipitation of driest month	0	0	0.1	0.4	0	0

Variables in bold were selected for the final model.

Table S3 Neutrality tests Fu's *F* and Tajima's *D* at the regional and supra-regional groupings.

	*trn*H-*psb*A				gapC-572				gapC-575			
Code	*D*	*P-value*	*F*	*P-value*	*D*	*P-value*	*F*	*P-value*	*D*	*P-value*	*F*	*P-value*
Andalusia	1.634	0.945	3.412	0.911	-1.732	**0.02**	-3.381	**0.002**	-2.104	**0.005**	-7.578	**<.001**
Azores	-1.411	0.071	-1.592	0.064	-2.103	**0.001**	-6.956	**<.001**	-1.608	0.188	-4.683	N.A.
Basque Country	0.35	0.688	0.462	0.64	-2.006	**0.003**	-5.317	**<.001**	-1.7334	**0.012**	-3.147	**0.002**
Canary Isl.	-0.271	0.421	0.703	0.632	-1.879	**0.008**	-5.036	**0.001**	-1.860	**0.008**	-7.928	**<.001**
Czech Republic	0	1	0	1	-1.491	**0.07**	-1.546	**0.018**	-2.121	**0.002**	-3.415	**0.003**
Galicia	0.735	0.802	1.279	0.789	-1.731	**0.015**	-2.9556	**<.001**	-1.9556	**0.004**	-4.434	**<.001**
Ire-Wal-Bri*	1.503	0.922	1.139	0.760	-1.881	**0.006**	-4.979	**<.001**	-2.349	**<.001**	-10.947	**<.001**
Italy	0	1	0	N.A.	-2.056	**0.003**	-5.655	**<.001**	-1.943	**0.004**	-3.698	**0.001**
Luxembourg	0	1	0	N.A	-1.706	**0.021**	-2.527	**0.004**	0	1	-0.879	0.078
Madeira	0.541	0.774	2.033	0.797	-1.159	0.155	-0.649	0.105	0	1	0	N.A.
Vosges du Nord	0.019	0.658	1.523	0.727	-1.562	**0.021**	-1.964	**0.008**	-1.697	**0.029**	-2.449	**0.001**
North	-0.224	0.412	-1.018	0.347	-2.366	**<.001**	-340	**<.001**	-2.478	**<.001**	-340	**<.001**
Cantabrian	0.388	0.68	0.667	0.655	-2.239	**<.001**	-11.044	**<.001**	-2.244	**<.001**	-10.426	**<.001**
South	0.447	0.713	1.768	0.802	-2.392	**<.001**	-23.066	**<.001**	-2.399	**<.001**	-28.768	**<.001**

The results for the regional level are shown above the dotted line, and those for the supra-regional level appear below the dotted line. *Ireland-Wales-Brittany. North, Northern evolutionary unit; Cantabrian, Cantabrian Cornice; South, Southern evolutionary unit. Statistically significant values are indicated in bold text. N.A., not applicable.

Methods S1 Expanded version of Materials and Methods.

DNA extraction, PCR amplification and sequencing

Total genomic DNA of the 312 sampled individuals was extracted from cleaned filaments of gametophytes, and from fresh, silica dried, or herbaria fronds of sporophytes using the NucleoSpin Plant II kit (Macherey-Nagel GmbH & Co. KG; Düren, Germany), following the manufacturer's instructions.

Plastid DNA for the intergenic spacer *trn*H-*psb*A was amplified by polymerase chain reaction (PCR) in all 309 individuals of *V. speciosa* and two outgroup species; we could not amplify this region in one of the two *V. birmanica* accessions. A previous survey using primers trnH^{GUG} (Tate & Simpson, 2003) and psbA (Sang *et al*., 1997) demonstrated the potential usefulness of this plastid region, inducing us to design specific primers for *V. speciosa* VS-trnH^{GUG}2 (5'- TGGATTCACAAATCCATTGC-3')

and VS-psbA2 (5'-CGTAATGCTCATAACTTCCCTCT-3'). PCR reactions were performed in 25-µL, under standard conditions (Innis *et al*., 1990). Automated sequencing of the purified PCR products was performed using VS-trnH^{GUG}2 primer on a 3100-Avant Genetic Analyzer (Applied Biosystems, Foster City, California, USA) in the "Centro de Instrumentación Científica" of the University of Granada (Spain).

We used the *gapCp* gene as nuclear marker, which has been used successfully in various phylogenetic studies of Hymenophyllaceae species, including the genus *Vandenboschia*, and has shown variation at infraspecific levels (e.g. Ebihara *et al*., 2005, 2009; Nitta *et al*., 2011). Due to the tetraploid nature of *V. speciosa* (Manton, 1950), only gametophytes (diploids) were used for nuclear marker amplification. In total, 150 *V. speciosa* gametophytes from 36 populations and one sporophyte of each *V. birmanica* accessions, and *V. davallioides* were analysed for the nuclear *gapCp* gene. PCR reaction conditions were the same than those for the plastid region except for the addition of 5 µL NZYTaq 5x Optimizer Solution (NZYTech, Lisbon, Portugal). The primers used for PCR amplification were GapC-7FA and GapC-BR-1 (Ebihara *et al*., 2005) and cycling parameters followed Ebihara *et al*. (2005). Purified PCR products were ligated into the pSC-A-amp/kan vector of the StrataClone PCR Cloning Kit (StrataGene, Agilent technologies, CA, USA) and cloned in StrataClone SoloPack competent cells (StrataGene), following the manufacturer's recommendations. Between three and five recombinant clones were sequenced per individual, using the generic primer M13F at the "Centro de Instrumentación Científica" of the University of Granada (Spain).

For both cpDNA and nuclear *gapCp*, sequences were edited and aligned, using the Clustal algorithm, in the alignment editor BIOEDIT v7.0.5.3 (Hall, 1999), and then adjusted by eye. All sequences were deposited in European Nucleotide Archive (accessions XXX–XXX).

Identification of gapCp *paralogues by phylogenetic analysis*

By visual inspection of the *gapCp* alignment, we detected a set of nucleotide positions that differentiate two types of sequences. To establish the homology relationship between these variants, we carried out a maximum likelihood (ML) phylogenetic analysis, including the *V. speciosa* sequences, those established by us for *V. birmanica* and *V. davallioi*des, and the *gapCp* sequences in GenBank for *V. radicans* group and *V. auriculata* (subgenus *Lacosteopsis*; as outgroup species) (Ebihara *et al*., 2005; accession

numbers: AB196370-AB196419). Inter-copy recombinant and single-point mutation sequences were detected with DNAsp v5.10 (Librado & Rozas, 2009) and removed from alignment (Schneider *et al*., 2013), which finally included 499 sequences in total. For phylogeny reconstruction, we used haplotypes instead all sequences; haplotypes were detected with ARLEQUIN v3.5.2.2 (Excoffier & Lischer, 2010). Phylogenetic analysis was conducted with PhyML v3.0 (Guindon *et al*., 2010) through the PhyML web server (http://www.atgc-montpellier.fr/phyml-sms/), with the nucleotide substitution model automatically selected by the Smart Model Selection tool integrated in the PhyML web server (Lefort *et al*., 2017), and the tree searching starting with five random trees obtained by BIONJ algorithm (Gascuel, 1997) and SPR as branch swapping method. Branch supports were assessed by the Shimodaira-Hasegawa-Like (SHL) implementation of the approximate likelihood-ratio test (Anisimova & Gascuel, 2006). Because the two *gapCp* variants were paralogues (called gapCp-572 and gapCp-575, see Results), we independently considered both copies in the subsequent analyses and only the largest intra-copy nonrecombining portion (detected with DNAsp).

Haplotype phylogeny and dating

Phylogenetic relationships among cpDNA haplotypes of *V. speciosa* and the outgroup species were inferred using Bayesian Inference (BI), with MrBayes v3.1.2 (Ronquist & Huelsenbeckn, 2003), and the best-fit nucleotide substitution model (HKY), as implemented in MrModeltest version 2.3 (Nylander, 2004) and considering the Akaike's information criterion (Akaike, 1973). The analysis was based on two million generations with four simultaneous runs (16 Markov chain Monte Carlo chains) starting from random trees that were sampled every 100 generations. TRACER v1.7 (Rambaut *et al*., 2018) was used to check the stationary of the runs and the convergence between runs. The initial 25% of the trees that resulted were discarded as burn-in, and the remaining trees were used to build 50% majority rule consensus trees.

To relate genetic differentiation found among cpDNA haplotypes to Neogene-Quaternary events, we estimated divergence times using BEAST2 package (Bouckaert *et al*., 2014). We followed a two-step strategy for tree calibration, because we could not establish specific dates for the few known fossils (the oldest considered to be from the Upper Miocene; see Bozukov, 2008). An initial analysis, to estimate the divergence time of *V. speciosa* lineage, was implemented with sequences of the *rbc*L gene (due to

the higher availability of sequences in the nucleotide database), all taken from GenBank database (accession numbers on the resulting tree; see Results). This included 32 *rbc*L sequences in total, one of *V. speciosa* (Dubuisson *et al*., 1997), 21 of other 14 *Vandenboschia* species, and 10 sequences of eight external species to *Vandenboschia* selected as representative of the main lineages of Trichomanoids [*Crepidomanes* (3 species), *Trichomanes ankersii*, *Abrodictyum elongatum*] and *Hymenophyllum* (3 species) according to Schuettpelz and Pryer (2009). Phylogeny was calibrated employing the dates proposed by Schuettpelz and Pryer (2009) by constricting the origins for Hymenophyllales (209.5-158.5 million years ago (Ma), mean: 185.1 Ma), *Hymenophyllum* (50.5-40.8 Ma, mean: 41.9 Ma; used by us as the most external taxon), all Trichomanoid genera (152.0-141.0, mean: 147.3 Ma), for the common ancestor of genera *Abrodictyum*, *Cephalomanes*, and *Trichomanes* (135.1 Ma), for the common ancestor of *Vandenboschia* and *Crepidomanes* (117.8 Ma), and for the common ancestor to *Crepidomanes* (95.1 Ma). The partitioned (for codon positions) .xml file was made up in BEAUti v2 (Bouckaert *et al*., 2014) by using a GTR+G+I substitution model, selected by jModelTest2 (Darriba *et al*., 2012), an uncorrelated lognormal relaxed-clock model (Drummond *et al*., 2006), after rejecting the strict molecular clock with the software BASEML of the PAML package (Yang, 2007), and a Calibrated Yule Model as the tree prior. BEAST v2.5.2 (Bouckaert *et al*., 2019) was launched with 50 million generations sampling one tree and parameters every 5,000 generations. TRACER v1.7 was used to check chain convergence and effective sampling size of the parameters. The maximum clade credibility tree (MCC) summarizing the estimated mean age and the 95% confidence intervals from post-burn-in (10%) trees was calculated with TREEANNOTATOR v2.5.2 (Bouckaert *et al*., 2014). Resulting MCC tree was visualized in FIGTREE v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

The second analysis concerned the *trn*H-*psb*A haplotypes of *V. speciosa*, using the sequences of eight species of *Vandenboschia* (six sequences taken from Genbank, accession numbers on the tree in Results, and two generated by us of *V. birmanica* and *V. davallioides*) as outgroup. Two calibration points were defined, using the dates from the first analysis. One for the common ancestor of the *V. davallioides* group (including *V. davallioides*, *V. kalamocarpa*, *V. nipponica*, *V. speciosa*, and *V. subclathrata*), and another for the common ancestor of the same taxa except *V. davallioides*. BEAST package was used following the same procedure described above, but using the HKY substitution model, and the Coalescent Constant Population Model as tree prior.

Demographic Analyses

The neutrality tests Fu's *F* (Fu, 1997) and Tajima's *D* (Tajima, 1989) were used to detect possible historical demographic processes (expansion or contraction), using ARLEQUIN. Both tests were performed considering geographical regions and evolutionary units, and with each molecular marker. The level of significance of both statistics was calculated by 10,000 simulated samples. In addition, cpDNA sequences were used to test for evidence of population size fluctuations within *V. speciosa* evolutionary units and also within Cantabrian range by constructing Bayesian Skyline Plots with BEAST (BSP; Drummond *et al*., 2005). Analyses were performed assuming a strict clock model with a fixed clock rate inferred from the infraspecific BEAST analysis described earlier (1.16×10^{-3} s/s/y), a coalescent Bayesian skyline prior as a tree model with five groups of coalescent intervals, and the HKY model with empirical frequencies and the transition-transversion parameter (kappa) fixed to 4.3 according the selected model. Also, we included a root height prior for each analysis, with a normal distribution, derived from the infraspecific dating analysis (mean 5.1 My, SD = 2.1). Analyses were run for 20 million generations and 2,000 as sample frequency. We used TRACER to examine the trace files for convergence and plots after discarding the burn-in (10%).

To reconstruct the historical migration routes in *V. speciosa* and locations of internal nodes of the haplotype phylogeny, cpDNA sequences were analysed using the marginal approximation of the structured coalescent implemented in BEAST2 package MASCOT v1.2.2 (Marginal Approximation of the Structured COalescenT; Müller *et al*., 2018). We defined 3 areas based on the evolutionary unit structuring: North, South, and Cantabrian. Prior substitution model was as performed in the BSP analysis. A lognormal distribution was used as prior for both the effective population size (Ne; M = 0 and S = 1) and the clock rate ($M = 1.16 \times 10^{-3}$, in the real space, S = 0.25), following the suggestion in "Taming the Beast" (https://taming-the-beast.org; Barido-Sottani *et al*., 2018). MCMC analysis was run for 50 million generations, sampling every 5,000 generations. The log file was analysed with TRACER to confirm adequate sample size and to determine the migration rates and Ne estimates. TREEANNOTATOR was used to summarize the trees in a MCC tree after descarding the first 10% as burn-in. FIGTREE was used to visualize the MCC tree.

References

Akaike H. 1973. Information theory and an extension of the maximum likelihood principle. In Petrov BN, Caski F, eds. *Proceedings of the Second International Symposium on Information Theory.* Budapest: Akademiai Kiado, 267–281.

Anisimova M, Gascuel O. 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic Biology* 55: 539–552.

Barido-Sottani J, Bošková V, Plessis LD, Kühnert D, Magnus C, Mitov V, Müller NF, Pečerska J, Rasmussen D, Zhang C et al. 2018. Taming the BEAST—A community teaching material resource for BEAST 2. *Systematic Biology* 67: 170–174.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* 10: e1003537.

Bozukov V. 2008. Genus *Trichomanes* in the European and Bulgarian Cenozoic macroflora. *Phytologia Balcanica* 14: 3–9.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9: 772.

Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* 4: e88.

Dubuisson JY. 1997. *rbc*L Sequences: A promising tool for the molecular systematics of the fern genus *Trichomanes* (Hymenophyllaceae)? *Molecular Phylogenetics and Evolution* 8: 128–138.

Ebihara A, Ishikawa H, Matsumoto S, Lin SJ, Iwatsuki K, Takamiya M, Watano Y, Ito M. 2005. Nuclear DNA, chloroplast DNA, and ploidy analysis clarified biological complexity of the *Vandenboschia radicans* complex (Hymenophyllaceae) in Japan and adjacent areas. *American Journal of Botany* 92: 1535–1547.

Ebihara A, Matsumoto S, Ito M. 2009. Hybridization involving independent gametophytes in the *Vandenboschia radican*s complex (Hymenophyllaceae): a

new perspective on the distribution of fern hybrids. *Molecular Ecology* 18: 4904–4911.

Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564–567.

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915–925.

Gascuel O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Molecular Biology and Evolution* 14: 685–695.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59: 307–321.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.

Innis MA, Gelfand DH, Snlnsky JJ, White TJ. 1990. PCR protocols: a guide to methods and applications. New York, USA: Academic Press.

Lefort V, Longueville JE, Gascuel O. 2017. SMS: smart model selection in PhyML. *Molecular biology and evolution* 34: 2422–2424.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.

Manton I. 1950. Problems of cytology and evolution in the Pteridophyta. Cambridge, UK: Cambridge University Press.

Müller NF, Rasmussen D, Stadler T. 2018. MASCOT: Parameter and state inference under the marginal structured coalescent approximation. *Bioinformatics* 34: 3843-3848.

Nitta JH, Ebihara A, Ito M. 2011. Reticulate evolution in the *Crepidomanes minutum* species complex (Hymenophyllaceae). *American Journal of Botany* 98: 1782–1800.

Nylander JAA. 2002. MrModeltest. Program distributed by the autor. Department of Systematic Biology, Uppsala University, Uppsala, Sweden.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* 67: 901–904.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

Sang T, Crawford DJ, Stuessy TF. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of Paeonia (Paeoniaceae). *American Journal of Botany* 84: 1120–1136.

Schuettpelz E, Pryer KM. 2009. Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. *Proceedings of the National Academy of Sciences* 106: 11200–11205.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.

Tate JA, Simpson BB. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany* 28: 723–738.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 24: 1586–1591.

Notes S1 Notes on the utility of the *gapCp* sequences for detect the allopolyploid origin of *Vandenboschia speciosa***.**

According to the chromosomal counts of Manton (1950) *V. speciosa* is a tetraploid species whose origin is not clear, although Ebihara has suggested that it could be an allotetraploid (unpublished data; suggested in Ebihara *et al*., 2007). Ní Dhúill (2014; unpubl. Ph.D) found variation in cytotypes (diploid, triploid and tetraploid sporophytes) in several Irish populations using flow cytometry, suggesting that the origin of the species could be more complicated than previously thought. Rumsey, considering the morphological and cytotype variation, raises the possibility that *V. speciosa* has had polytopic origins and that there may be more than one cryptic taxon (pers. com.).

However, the monophyly of the cpDNA haplotypes suggests the rejection of the hypothesis of species' polytopic origins and cryptic taxa; unless the maternal parent has always been the same taxon in each of the allopolyploidization processes (and then we are seeing only the haplotypes of the maternal species) or the parental taxa were sister taxa (monophyly due to sharing of a most recent common ancestor; although the geographical structure of the haplotypes argues against this possibility). Nuclear biparentally inherited markers, as the *gapCp* gene, could help reveal the origin of the species.

The two types of paralogues we detected for the *gapCp* gene, with differential additivity to the *gapCp* of species of well-differentiated clade (*V. nipponica*-*V. kalamocarpa* vs. *V. birmanica*, Fig. S1) suggest the allopolyploid origin of *V. speciosa*. However, the presence of both paralogues in the external species used (especially in *V. davallioides*) prevents us from concluding that it is an allotetraploid species, and that both copies are in *V. speciosa* by inheritance from ancestral polymorphic lineages. The utility of *gapCp* in resolving polyploid origin in ferns, including *Vandenboschia* species, has been evidenced by several authors (Ebihara *et al*., 2005, 2009a, 2009b; Schuettpelz *et al*., 2008; Nitta *et al*., 2011), but some of these studies also warn of some disadvantages as the detection of several intraspecific copies or of multiple alleles in single specimens (Schuettpelz *et al*., 2008; Nitta *et al*., 2011; Rothfels *et al*., 2013). Nitta *et al*. (2011), studying the complex *Crepidomanes minutum* (Hymenophyllaceae) detected multiple *gapCp* alleles from single specimens, especially within one of the clades recovered by their phylogenetic hypothesis. These authors ruled out that these were paralogous genes because of all *C. minutum* species complex *gapCp* sequences were monophyletic, and were easily aligned to each other and the outgroup (Nitta *et al*., 2011). As we have indicated above, the presence of both copies in the external species of our analysis and the fact that these sequences intermingle with those determined from *V. speciosa* imply a relationship of paralogy between the copies detected, pointing to a phenomenon of gene duplication that occurred less before the diversification of the ancestral lineage and which resulted in the groups including *V. speciosa* and *V. davallioides*. Duplications of this gene have also been observed in the genera *Hymenophyllum* (Hymenophyllaceae) and *Dipteris* (Dipteridaceae) (Rothfels *et al*., 2013).

References

Ebihara A, Ishikawa H, Matsumoto S, Lin SJ, Iwatsuki K, Takamiya M, Watano Y, Ito M. 2005. Nuclear DNA, chloroplast DNA, and ploidy analysis clarified biological complexity of the *Vandenboschia radicans* complex (Hymenophyllaceae) in Japan and adjacent areas. *American Journal of Botany* 92: 1535–1547.

Ebihara A, Iwatsuki K, Ito M, Hennequin S, Dubuisson JY. 2007. A global molecular phylogeny of the fern genus *Trichomanes* (Hymenophyllaceae) with special reference to stem anatomy. *Botanical Journal of the Linnean Society* 155: 1–27.

Ebihara A, Matsumoto S, Ito M. 2009a. Hybridization involving independent gametophytes in the *Vandenboschia radican*s complex (Hymenophyllaceae): a new perspective on the distribution of fern hybrids. *Molecular Ecology* 18: 4904–4911.

Ebihara A, Matsumoto S, Ito M. 2009b. Taxonomy of the reticulate *Vandenboschia radicans* Complex (Hymenophyllaceae) in Japan. Acta Phytotaxonomica et Geobotanica 60: 26–40.

Manton I. 1950. Problems of cytology and evolution in the Pteridophyta. Cambridge, UK: Cambridge University Press.

Nitta JH, Ebihara A, Ito M. 2011. Reticulate evolution in the *Crepidomanes minutum* species complex (Hymenophyllaceae). *American Journal of Botany* 98: 1782–1800.

Rothfels CJ, Larsson A, Li FW, Sigel EM, Huiet L, Burge DO, Ruhsam M, Graham SW, Stevenson DW, Wong GK *et al*. 2013. Transcriptome-mining for single-copy nuclear markers in ferns. *PLoS One* 8: e76957.

Schuettpelz E, Grusz AL, Windham MD, Pryer KM. 2008. The utility of nuclear gapCp in resolving polyploid fern origins. *Systematic Botany* 33: 621–629.

Chapter 3. Genetic diversity and phylogeography of the relict fern *Culcita macrocarpa*: influence of clonality and reproduction system on genetic structure

Samira Ben-Menni Schuler, Jim Provan, Ana Teresa Romero García, Víctor N. Suárez Santiago

Manuscript in preparation

1. Introduction

In plant phylogeography, there are very few studies addressing the history of ancient taxa, as those of the European lauroid forest (e.g. Rodríguez-Sánchez et al., 2009; García-Verdugo et al., 2013; Chen et al., 2014). The most studies are centred in the effects of Pleistocene glaciations on the plant distributions (Petit et al., 2005). Thus, in Europe, plant phylogeography has allowed to identify the main routes of postglacial colonization, glacial refugia and suture zones (e.g. Hewitt, 1996, 1999; Comes and Kadereit, 1998; Taberlet, 1998; Cruzan and Templeton, 2000; Stewart and Lister, 2001; Petit et al., 2003; Hampe and Petit, 2005; Gómez and Lunt, 2006; Provan and Bennett, 2008; Médail and Diadema, 2009), but there are very few phylogeographic data that complete or contrast the hypotheses proposed by historical biogeographic studies on the history of ancestral taxa and the factors responsible for their demise and current distribution (as in Mai, 1989, 1991; Barrón, 2003).

This lack of phylogeographic data is especially important among ferns (e.g. Trewick et al., 2002; Hunt et al., 2009; Jiménez et al., 2010; Bystriakova et al., 2014, Maccagni et al., 2017), the second most diversified terrestrial plant group. Numerous ferns are considered lineages of Tertiary origin, especially most of the Mediterranean region (Pichi-Sermolli, 1979, 1991), which were the main component of the herbaceous layer of the European lauroid forest (Benito Garzón and Sainz Ollero, 2002; Barrón, 2003; Barrón & Peyrot, 2006), and which survived the Quaternary glaciations in various areas (Iberian Peninsula, Italy, Balkans, and Greece) that served as glacial refugia (Vogel et al., 1999a). Thus, the origin and initial stages of diversification of these species can be associated with powerfull geological and climatic processes before the Pleistocene; and even they may be the initial responsible for the current geographical distribution of these ferns.

On the other hand, the fern biological peculiarities (in relation to spermatophytes) make them interesting species to explore the influence of these peculiarities on the population processes responsible for genetic structure. The ferns are organisms with high dispersive capacity (Wolf et al., 2001), by spores, which favors population connectivity and decreased genetic structure (Soltis & Soltis, 1989); however, and especially for the European Tertiary ferns, the scarcity and the disjunct distribution of suitables habitats with high environmental humidity and warm temperatures can result in strong genetic

structuring (e.g. Trewick et al., 2002; Shepherd et al., 2008; Hunt et al., 2009; Jiménez et al., 2010; Wang et al., 2011; Bystriakova et al., 2014; Maccagni et al., 2017). In ferns fertilization is always a post-dispersal process (it occurs on the gametophyte originated from spores) and then the probability of colonization of a new habitat, and as a consequence the population genetic composition, will depends of the breeding system of the species; less probable for ferns with outcrossing and inter-gametophytic selfing, resulting in more diverse populations, while those with intra-gametophitic selfing are able to establish a new population even from just one spore (Suter et al., 2000; de Groot et al., 2012a). Moreover, many ferns also have the ability to propagate vegetatively, what affect the quantity and distribution of genetic diversity. Clonal reproduction will impoverish genetically the populations by increasing the self-fertilization rate (Carrillo-Angeles et al., 2011) or because interclonal competition, and elimination of less adapted clones (even leading to monoclonal populations), occurs (Silvertown, 2008). However, clonality tends to increase the heterozygosity by accumulation of mutation and divergence between alleles (Balloux et al., 2003; de Meeûs et al., 2007; although see Douhovnikoff & Leventhal, 2016). Population differentiation is drastically reduced when reproduction tends towards strict clonality, due to the genetic diversity is maintained within clonal lineages (Balloux et al., 2003).

In this study we focus on *Culcita macrocarpa* C. Presl (Culcitaceae), a diploid tree fern considered to be a relict from the Palaeotropical geoflora (Pichi-Sermolli, 1979). It is an Iberian-Macaronesian endemism (Andalusia, Cantabrian Cornice, and archipelagos of Azores, Canary Islands and Madeira), and the only European representative of the genus, since the other known species (*Culcita coniifolia*) has an American distribution. It lives restricted to shade places with mild temperatures throughout the year and high atmospheric and edaphic humidity; conditions found in valleys with strong slopes and near the coast, usually at altitudes below 300 m, or in higher areas asociated with fog belts (even above 1,000 m), especially in Azores. The fragility of the habitats where *C. macrocarpa* is distributed and its disjunct distribution have led to be protected under the Bern Convention and the Habitats Directive (Anonymous, 1979, 1992). It is considered an endangered species in the Red List of Spanish Vascular Flora (Bañares et al., 2008).

From a reproduction point of view, culture experiments have shown that gametophytes of *C. macrocarpa* are initially male and later hermaphrodite (Stokey, 1930, Quintanilla et al., 2005), what seems to favor intragametophytic selfing (Klekowski and Lloyd,

1968). Quintanilla et al. (2007) studied the genetic variation of six populations from Galicia (northwestern Iberian Peninsula) using 13 enzymatic systems. Among other things, these authors wanted to assess the concordance between genotypic frequencies and the breeding system. However, the total lack of variation found (one only mulilocus genotype shared by all individuals and populations) prevented these authors to test for random mating, and what did show was the intense clonal reproduction that the species presents (Quintanilla et al., 2007). *Culcita macrocarpa* has a creeping rhizome that can reach more than one meter in length, bifurcating and giving rise new shoot apices that form adventitious roots; each shoot apex will constitute a ramet (Quintanilla et al., 2000, 2005, 2007). The lack of genetic variation was attributed by the authors to genetic drift associated with the reduction in the number of populations during the last ice age and subsequent founder effects during Holocene expansion (Quintanilla et al., 2007).

Despite being an emblematic fern in Europe, nothing is known about its origin and evolutionary history, nor tested its reproductive system with resolutive molecular markers and evaluated the effect of clonality on its intraspecific genetic structure. In order to quantify the genetic variation and its distribution within and between populations, and to know the habitat suitability for the species in the past and the future, we provide genetic data from microsatellite nuclear loci, one plastid marker, and species distribution modelling (SDM) trying to cover the following objectives: 1) To infer the phylogeographic history. The biogeographic pattern of the species, typical of a relictual taxon, makes *C. macrocarpa* a good system to assess the impact of both pre-Pleistocene and Quaternary geological-climatic events on population processes that generate genetic structure, and how could affect the future climate change in this evolutionary pattern. Phylogeographic data from ancient taxa as *C. macrocarpa* are important to know about the demise of the Palaeotropical geoflora in Europe, and in making broad biogeographic staments in general. 2) To estimate the importance of sexual and vegetative reproduction on population composition and to evaluate the effect of clonality on its intraspecific genetic structure.

2. Material and Methods

2.1. Plant Material

A total of 443 individuals of *C. macrocarpa* were obtained from 17 populations in 4 geographical regions across its distribution range: Andalusia, Azores, Canary Islands and the Cantabrian Cornice. In Madeira, where the species is also present, we could not find any population. The number of populations per region was between one and five, and the number of sampled individuals per population varied from 6 to 30 (Table 1; Fig. 1).

Table 1. Sampling details of *C. macrocarpa* populations used in the present study.

Code	Location	Voucher	Geographical coordinates	Sample size	
				Microsatellites	cpDNA
Andalusia					
ALM	Cádiz: Almoraima		N36.304°/W5.520°	6	3
CRM	Cádiz: Cabecera del río de la Miel		N36.105°/W5.528°	30	5
PIN	Cádiz: Laja del Pinarejo		N36.188°/W5.589°	30	5
RM	Cádiz: Río de la Miel		N36.112°/W5.507°	29	5
SDN	Cádiz: Sierra del Niño		N36.186°/W5.610°	30	5
Azores					
CAR	Terceira: Algar do Carvão	GDA63533	N38.727°/W27.215°	30	5
CID	São Miguel: Sete Cidades	GDA63534	N37.835°/W25.788°	29	5
FOG	São Miguel: Lagoa do Fogo	GDA63532		30	5
NAT	Terceira: Gruta do Natal	GDA63531	N38.738°/W27.264°	30	5
Canary Isl.					
IJU	Tenerife: Ijuana	GDA63536	N28.560°/W16.172°	30	4
Cantabrian Cornice					
BER	Bizkaia: Bermeo	GDA63539	N43.392°/W2.734°	12	5
BAK	Bizkaia: Bakio		N43.425°/W2.845°	18	5
CUN	Asturias: San Esteban de Cuñaba	GDA63537	N43.277°/W4.676°	30	5
EUM	A Coruña: Eume	GDA63535	N43.404°/W8.087°	30	5
LIE	Cantabria: Liendo		N43.375°/W3.383°	30	5
NUE	Asturias: Nueva de Llanes	GDA63538	N43.421°/W4.954°	30	5
SEI	A Coruña: Seixo	GDA63530	N43.706°/W7.946°	19	5

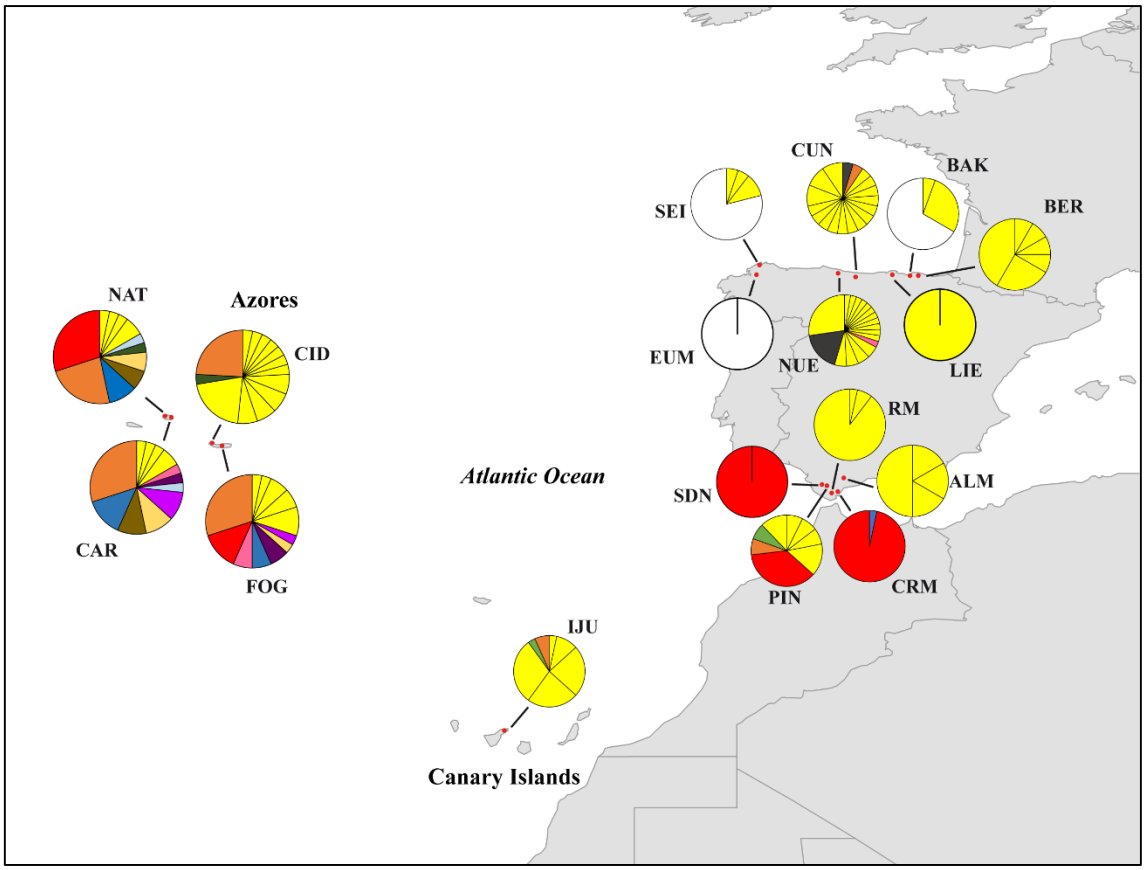


Figure 1. Map of the *Culcita macrocarpa* sampling locations. Frequency of each multilocus lineage (MLL) based on microsatellite data is represented as pie charts per location. Colours represent the different MLLs that are shared among sites, and yellow represents private MLLs that only appear in one site. See Table 1 for full name of locations. Note that pie chart size is the same for all locations and does not represent the number of individuals.

2.2 DNA Extraction, Microsatellite Genotyping, and cpDNA Sequencing

Total genomic DNA of the 443 individuals was extracted from silica dryed fronds following a modified low-salt CTAB extraction protocol (Arseneau et al., 2017). All individuals were genotyped for eight recently developed microsatellite loci described in Ben-Menni Schuler et al. (in prep.), and following the PCR conditions explained in that study. Genotyping was peformed on an ABI PRISM® 3100-Avant Genetic Analyzer (Applied Biosystems, Foster City, CA, United States). Alleles were scored using GENEMARKER v1.85 (SoftGenetics, State College, PA, United States).

For ptDNA analysis, a subsample of 82 individuals (3-5 per population) was used. The plastid marker used was the intergenic spacer *rps4-trnL*, the only polimorphic region of the 13 tested (introns in *rpl16*, *rps16*, *trnG*^{UCC}, *trnL*, and *ycf3*, and the intergenic spacers, *rpl32-trnL*, *rps4-trnL*, *rps16-matK*, *trnD-psbM*, *trnD-rpoB*, *trnH-psbA*, *trnL-*

trnF, and *trnS*^{GGA}-*trnG*^{UCC}). All primer pairs used for PCR amplification were designed from the partial chloroplast genome of *Plagiogyria glauca* (GenBank accession number: KP136831; Supl. Table 1). PCR reactions were performed in 25-µL reactions containing 50 ng of genomic DNA, 1.25 µM of each primer, 2 mM $MgCl_2$, 1.5 mM dNTPs, 2.5 µL Biotools 10× Reaction Buffer, and 1.5 units of Biotools DNA polymerase (Biotools B&M Laboratories S.A., Madrid, Spain). Cycling parameters consisted of 5 min of denaturing at 94°C; followed by 35 cycles of 94°C for 1 min, 50°C for 1 min, and 72°C for 1 min; and a final extension at 72°C for 10 min. Sequencing was performed on an ABI PRISM® 3100-Avant Genetic Analyzer (Applied Biosystems, Foster City, CA, United States). The resulting sequences were aligned by eye in the alignment editor BIOEDIT v7.0.5.3 (Hall, 1999).

2.3 Clonality and Genetic diversity

2.3.1 Microsatellites

In order to infer the clonal identity of the sampling units (all individuals sampled), first we tested the resolutive power of the eight microsatellite markers by estimating the genotype accumulation curve using the function *genotype_curve* from the R package POPPR v2.8.3 (Kamvar et al., 2014). This function sample, using a Monte Carlo procedure, random subsets of loci and examine the robustness of the inferred clonal memberships. Then, we used MLGsim v2.0 (Stenberg, Lundmark & Saura, 2003) for calculating the probability that repeated multilocus genotypes (MLGs) originate from different sexual reproduction events (p_{sex}; being different genets), based on the observed allele frequencies and the sample size of the data set, and taking into account departures from Hardy-Weinberg equilibrium (HWE) when estimating genotypic probabilities (p_{gen} (F_{IS}), for a more conservative estimate of p_{sex}; Arnaud-Haond et al., 2007). The significance values of p_{sex} were obtained by comparison with the distribution of 1,000 simulated *p*sex values. Finally, to define the clonal lineages or multiple locus lineages (MLLs; i.e. different MLGs belonging to a distinct or the same clone) we analized the distribution of the frequencies of genetic distances between pairs of MLGs, with the function *mlg.filter* and using Bruvo distances on POPPR. The genetic threshold distance under which two MLGs were considered the same MLL was estimated using the farthest neighbor method.

The clonality descriptors were calculated with the function *poppr* on POPPR as follows: a) to characterize the clonal richness we calculated: number of MLLs, number of expected MLLs (eMLLs), and clonal richness (R) corrected for sampling size; b) to characterize the genotype diversity we calculated: the Simpson's index (lambda; corrected for sampling size), and the clonal eveness index (E.5), which shows how equally each MLL is represented. Finally, we calculated the standarized association index (r_d; Agapow & Burt, 2001) to test the predominant reproductive model (sexual, where linkage among loci is not expected, *vs*. clonal, where significant disequilibrium is expected due to linkage among loci). This latter index was also estimated applying correction for clones, and then using only one individual per MLL, in order to test the effect of partial clonality. The significance of r_d was tested with a permutation test (10,000 permutations). All descriptors were estimated both at population and geographical region levels.

To calculate genetic diversity descriptors we used all sampling units, following the recommendation of De Meeûs et al. (2006) and because we consider that this way the real genetic structure of *C. macrocarpa* populations is better represented. Thus, we calculated: the number of alleles (*A*) and allelic richness (*Ar*), rarefacting to the smallest sample size, using the R package HIERFSTAT v0.04-30 (Goudet, 2005) with the functions *allele.count* and *allelic.richness*, respectively; the observed and expected heterocigosity (H_O and H_E, respectively), and fixation index F_{IS} and HWE using GENODIVE v3.0 (Meirmans & Van Tienderen, 2004). Furthermore, in order to test the effect of asexual reproduction on intrapopulation genetic diversity, we also calculated genetic descriptors (*Ar*, H_O, H_E and F_{IS}) using only one individual per MLL per population, and Student's *t* tests were performed to explore significant differences between data sets (including vs. excluding clones). All genetic descriptors were calculated both at population and geographical region levels.

2.3.2 ptDNA

Genetic diversity was assessed by the number of haplotypes (*ha*), haplotype diversity (*Hd*), and nucleotide diversity (π) calculated using ARLEQUIN v3.5.2.2 (Excoffier & Lischer, 2010). All diversity indices were calculated both at population and geographical region levels.

2.4 Genetic structure and phylogeography

Standard and hierarchical analysis of molecular variance (AMOVA; Excoffier, Smouse & Quattro, 1992) were used to test for partitioning of genetic variability within samples, within and between populations, and between the 4 geographical regions. For microsatellites, these analyses were made with all sampling units and with only one individual per MLL per population, using the function *poppr.amova* on POPPR and the function *randtest*, running 1,000 replicates, for testing for significance. For cpDNA we used the program ARLEQUIN, and the significance was tested with 10,000 permutations. In addition, one more AMOVA analysis was carried out for both microsatellites and cpDNA considering the geographical distribution of the haplotypes and the result of the network analysis (see Results). Thus, we considered a fourth level of population grouping called supraregional grouping. Two supraregional groups were considered: South (Andalusia, Azores and Canary Islands) and the Cantabrian Cornice. Population genetic structure was analysed using different approaches with our microsatellite data. First, pairwise F_{ST} values were calculated, with both all sampling units and with only one individual per MLL per population, between populations using GENODIVE; the significance of F_{ST} was tested by a permutation test with 10,000 permutations. We compared the values obtained with and without clones using the Spearman's correlation coefficient. Second, the Bayesian algorithm implemented in STRUCTURE v2.3.4 (Falush et al., 2003) was used to evaluate the number of genetic clusters (*K*) with both all sampling units and with clone correction. The number of clusters tested ranged from one to 18, with 10 replicates per *K*, using the no admixture model and independent allele frequencies. The burn-in period and Markov Chain Monte Carlo (MCMC) iterations were set to 50,000 and 10^6, respectively. The optimal number of clusters was estimated with the online tool STRUCTURESELECTOR (Li and Liu, 2018). We identified the uppermost hierarchical level of genetic structure using the delta *K*-method (ΔK; Evanno et al., 2005). To explore other levels of genetic partitioning, we used the four independent estimators proposed by Puechmaille (2016; MedMedK, MedMeaK, MaxMedK, and MaxMeaK) considering a membership coefficient threshold of 0.5. To align and visualize the STRUCTURE output across the 10 replicates, we used the online tool CLUMPAK (Kopelman et al., 2015). Third, the genetic structure was also assessed using a model-free multivariate statistics-based

clustering method, a discriminant analysis of principal components (DAPC) on R package ADEGENET (Jombart et al., 2010) using all sampling units. The function *xvalDapc* from ADEGENET was used to select by crossvalidation the correct number of principal components with 1,000 replicates using a training set of 90% of the data. The number of principal components was chosen based on the criteria that it had to produce the highest average percentage of successful reassignment and lowest root mean squared error (Jombart et al., 2010).

In order to explore the evolutionary relationships and geographical distribution of ptDNA haplotypes an haplotype network was reconstructed following the statistical parsimony method (Templeton et al., 1992) as implemented in TCS v1.21 (Clement et al., 2000).

To relate genetic differentiation found among cpDNA haplotypes to Neogene-Quaternary events, we estimated the haplotype divergence times using BEAST2 package (Bouckaert et al., 2014). We used the *rps4-trnL* sequence of *Plagyogyria glauca* obtained from the partial chloroplast genome as ougroup. Phylogeny was calibrated employing the dates proposed by Testo and Sundue (2016) by constricting the most recent common ancestor between the genera *Plagyogyria* and *Culcita* (116.61-96.94 million years ago (Ma), mean: 99.38 Ma). The .xml file was made up in BEAUti v2 (Bouckaert et al., 2019) by using a Jukes Cantor substitution model, selected by jModelTest2 (Darriba et al., 2012), an strict molecular clock model (Drummond et al., 2006), and a Coalescent Constant Population Model as a tree prior. BEAST v2.5.2 (Bouckaert et al., 2019) was launched with 50 million generations sampling one tree and parameters every 5,000 generations. TRACER v1.7 was used to check chain convergence and effective sampling size of the parameters. The maximum clade credibility tree (MCC) summarizing the estimated mean age and the 95% confidence intervals from post-burn-in (10%) trees was calculated with TREEANNOTATOR v2.5.2 (Bouckaert et al., 2019). Resulting MCC tree was visualized in FIGTREE v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

2.5 Gene Flow using microsatellite data

We tested the connectivity among populations by estimating the migration rates among them, in all cases with all sampling units. Thus, to know whether there was recent (over

two to three generations) gene flow between the populations, we estimated migration rates (*m*) between all individual populations using a Bayesian assignment test with the software BAYESASS v1.3 (Wilson & Rannala, 2003). As program settings, the default values were used (MCMC iterations, 3×10^6; length of the burn-in, 999,999; sampling frequency, 2,000; delta value, 0.15). Isolation by distance (IBD) was tested for the 17 populations using regression of pairwise F_{ST} distances [determined with GENODIVE using them transformed as $F_{ST}/(1 - F_{ST})$] and logarithms of geographical distances between populations, using a Mantel test in GENODIVE.

2.6 Species Distribution Modelling

To identify potential refugial and future distribution areas for *C. macrocarpa*, species distribution modelling (SDM) was performed. This analysis requires presence occurrence data of the studying species and environmental variables. As environmental data we used 19 BIOCLIM variables at a resolution of 2.5 arc-minutes (ca. 5 km) representing different time periods during past, present and future climatic conditions. Past and current climate data were available from the WorldClim database (www.worldclim.org; Hijmans et al., 2005) and included data for the current-day period (1,950– 2,000), the Last Glacial Maximum (LGM; *c.* 21 ka) simulated by CCSM model (the Community Climate System Model), and for the Last Interglacial period (LIG; *c.* 120 ka). We obtained predictions for future climatic conditions in 2.080 for the most impacting IPCC's climate scenario: RCP8.5 (Van Vuuren et al., 2011) available through the CCAFS Climate portal (www.ccafs-climate.org). Soil data were obtained from SoilGrids.org (Hengl et al. 2014) but was not used with past climatic conditions in LGM because of the lack of this type of maps. Highly correlated variables (Pearson's R ≥ 0.8) were reduced to eight uncorrelated variables (Table S2) used as predictors to calibrate the distribution models. Species occurrence data is a collection of references in databases (the Global Biodiversity Information Facility data portal (http://www.gbif.org/), the Biodiversity databank of the Canary Islands (http://www.biodiversidadcanarias.es/atlantis/common/index.jsf), and the Azores Biodiversity databank (http://www.atlantis.angra.uac.pt/atlantis/common/index.jsf), literature (Guitián, 2010; Romero et al., 2004; Soñora, 1992; Quintanilla & Amigo, 2001), and our own field records. A total of 379 presence records were finally compiled

(Figure S1). To perform the SDM we applied Maximum Entropy Modelling implemented in the software package MAXENT 3.4.1 (Phillips et al., 2006). Models were generated using cross-validation of 5 replicate runs. Model performance was assess based on the area under the receiver operating characteristic curve (AUC). The contribution of each predictor variable in the model was analysed by the permutation importance and percent contribution coefficients (Table S2). A final reduced model including the most important variables (Martínez et al., 2012), this is, Mean Diurnal Range and Minimum Temperature of Coldest Month, was finally computed.

3. Results

3.1. Clonality and genetic diversity

3.1.1. Microsatellites

A total of 120 different MLGs were detected between the 443 sampling units. The genotype accumulation curve showed that our eight microsatellite loci have a high power to discriminate between the MLGs of *C. macrocarpa*, since with seven loci almost 100% of the MLGs are resolved (Fig. S2). Only for one repeated MLG cannot be discarded independent origin by sexual reproduction (p_{sex} = 0.96, p = 0.115). This MLG was shared between two sampling units of different populations, CAR (Azores) and NUE (Cantabrian Cornise), and therefore, these and all sampling units of different populations with shared MLGs were maintained in the data set when clone correction was applied. The genetic threshold distance under which two MLGs were considered belong to the same MLL was 0.0391 (Fig. S3). After the MLGs were collapsed in MLLs, the total number of MLLs was 104 distributed among 130 individuals (genets) across the populations, with different number of clones between populations and geographical regions (Figure 1, Fig. S4a,b; Table 2). Between 1 and 23 MLLs were detected across all the populations. Thirteen MLLs were shared among populations, of which the most frequent showed a differential distribution between Andalusia, Azores and the Cantabrian Cornice (Fig. 1, MLLs: red, orange, and white and black; respectively). The rest of MLLs were private to the populations (Fig. 1; Table 2). The Cantabrian Cornice was the region that retained the highest number of MLLs (53) and clonal richness (R = 0.31), although this is not similarly distributed among the populations. Many Cantabrian populations had very low number of MLLs (even two are

monoclonal, EUM and LIE), while CUN (23 MLG; R = 0.759) and NUE (18 MLG; R = 0.586) harboured the highest values of all sampled populations. On the contrary, in Azores, although had fewer MLLs (35, R= 0.29), their populations harboured relatively high and similar levels of clonal richness (Table 2); and the expected MLL values (eMLL; Table 2) were even higher than those of the Cantabrian Cornice. The differential prevalence of clones between Azores populations and between those of the Cantabrian Cornice is reflected in the regional evenness index (Table 2; Fig. S4a,b). Although the CUN and NUE populations showed a proportionate distribution of clones, and therefore high evenness, the dominance of a few clones in other populations (except BER) made the Cantabrian Cornice the region with the lowest value of evenness. Azores, meanwhile, was the region with the highest proportionality in the distribution of clones (excluding the Canary Islands) and also with the greatest clonal diversity (lambda; Table 2). Conversely, Andalusia was the region with lowest values of clonal richness and genotype diversity; since the populations showed few MLLs and only some of them were dominant (Fig. 1, S4a,b; Table 2). In Canary Island, although the only known population presented moderate-low values of clonal richness, the genotype diversity indexes were close to those obtained in Azores and the most diverse populations of the Cantabrian Cornice (Table 2). All the populations for which the association index could be calculated, except FOG, NAT and IJU, and geographical regions, except the Canary Islands, showed a significant linkage disequilibrium when all sampling units were included. When only one individual was considered per MLL per population, only PIN, at the population level, and Andalusia and Cantabrian Cornice, at the regional level, presented significant disequilibrium (Table 2).

Table 2. Clonality descriptors in the populations of *C. macrocarpa* studied. Descriptors were separated into clonal richness, genotype diversity, and linkeage disequilibrium. N, number of individuals sampled; MLL, number of different multilocus lineages, or clonal lineages, with exclusive MLLs in brackets; eMLL, number of expected MLLs at the smallest sample size ≥ 6 based on rarefaction (Hurlbert, 1971); R, clonal richness (Dorken & Eckert, 2001); lambda, Simpson's index (Simpson, 1949); E.5, evenness (Pielou, 1975; Ludwig & Reynolds, 1988; Grünwald et al., 2003); r_d, standardized index of association (Agapow & Burt, 2001); r_d wc, r_d calculated considering only one individual per MLL per population.

		Clonal richness			Genotype diversity		Linkeage disequilibrium	
Population	**N**	**MLL (private)**	**eMLL**	**R**	**lambda**	**E.5**	**r_d**	**r_d wc**
Andalusia	125	16 (12)	6.56	0.121	0.579	0.464	0.402*	0.186*
ALM	6	4 (4)	4	0.6	0.8	0.812	NA	NA
CRM	30	2	1.33	0.034	0.067	0.438	NA	NA
PIN	30	8 (5)	4.12	0.241	0.623	0.528	0.728*	0.193*
RM	29	3 (3)	1.92	0.071	0.197	0.48	0.141*	0.5
SDN	30	1	1	0	0	NA	NA	NA
Azores	119	35 (25)	16.01	0.288	0.904	0.5	0.041*	-0.05
CAR	30	12 (4)	6.53	0.379	0.881	0.727	0.062*	-0.047
CID	29	14 (12)	6.98	0.464	0.901	0.719	0.140*	0.003
FOG	30	12 (5)	6.68	0.379	0.885	0.725	0.029	-0.063
NAT	30	11 (4)	6.03	0.345	0.855	0.713	0.005	-0.067
Canary Isl.	30	7 (5)	7	0.207	0.811	0.836	-0.018	-0.248
IJU	30	7 (5)	4.86	0.207	0.811	0.836	-0.018	-0.248
Cantabrian Cornice	169	53 (44)	14.93	0.31	0.853	0.347	0.185*	0.085*
BER	12	6 (6)	5.33	0.454	0.803	0.762	0.3*	-0.042
BAK	18	3 (2)	2.55	0.118	0.503	0.757	0.536*	0
CUN	30	23 (16)	9.2	0.759	0.982	0.913	0.04*	0.011
EUM	30	1	1	0	0	NA	NA	NA
LIE	30	1 (1)	1	0	0	NA	NA	NA
NUE	30	18 (16)	8.01	0.586	0.942	0.76	0.07*	0.024
SEI	19	4 (3)	2.84	0.166	0.38	0.52	0.725*	0.12
Total	443 (130)	104 (91)	7.69	0.292	0.926	0.384	0.288	

* $P < 0.05$; NA, not applicable; wc, without clones.

In total, 37 alleles were observed from the eight loci surveyed. Between 2 and 5 alleles per locus were detected across all the populations (Table 3). Allelic richness and the expected heterocigosity (H_E) significantly changed when only one individual per MLL per population was considered ($t = 3.05$; $p < 0.05$ and $t = -2.69$; $p < 0.05$ respectively), but the observed heterocigosity (H_O) and the fixation index (F_{IS}) not ($t = -1.41$; $p < 0.17$ and $t = -1.04$; $p < 0.31$ respectively). All populations, except ALM, deviated from HWE when all individuals were included, resulting in significantly negative values of F_{IS} in the populations CRM, RM, SDN, IJU, CUN and LIE. After consider only one individual per MLL per population F_{IS} values remained negative for these populations (except for SDN and LIE with only one MLL each) (Table 3). The negative values of F_{IS} were determined mainly by the locus *CM-AT19*, which showed fixed heterozygosity (for only two alleles) in almost all populations, and the high amount of monomorphic

loci in the different populations (Table S3). After exclude *CM-AT19*, only populations RM, IJU, and CUN showed negative F_{IS} (significant only in RM; data not shown). Considering the regions, the F_{IS} values showed a significant excess of homozygotes (except Canary Island, with only one population), both including all sampling units and only one individual per MLL per population. At the population level, it was the Azorean and Cantabrian (especially BER, NUE, and CUN) populations that showed the highest diversity values and the Andalusians the least; however, at regional level Andalusia showed similar levels of diversity than the other regions (Table 3).

Table 3. Genetic diversity indices for microsatellites and cpDNA sequences in the populations of *C. macrocarpa* studied. Indices were calculated including all individuals sampled per population and including only one individual per MLLs per population (wc; without clones) N, number of individuals sampled; *A*, number of alleles with unique alleles in brackets; *Ar*, allelic richness at the smallest sample size (12 and 2 for populations with clones and without clones, respectively; 60 and 12 for geographical regions with clones and without clones, respectively) based on rarefaction; H_O, observed heterozygosity; H_E, expected heterozygosity (Nei, 1987); F_{IS}, inbreeding coefficient (Weir & Cockerham, 1984); *ha*: number of haplotypes with unique haplotypes in brackets; *Hd*: haplotype diversity; π: nucleotide diversity ($x10^2$).

		Microsatellites									cpDNA		
Population	N	*A* (private)	*Ar*	*Ar* wc	H_O	H_O wc	H_E	H_E wc	F_{IS}	F_{IS} wc	*ha*	*Hd*	π
Andalusia	125	18 (3)	2.228	2.187	0.164	0.215	0.331	0.426	0.505*	0.494*	3 (1)	0.38	0.17
ALM	6	11 (1)	1.375	1.165	0.146	0.156	0.146	0.167	0.000	0.063	1	0.00	0.00
CRM	30	10	1.15	1.146	0.129	0.188	0.067	0.125	-0.938*	-0.500	1	0.00	0.00
PIN	30	12	1.498	1.255	0.163	0.250	0.232	0.267	0.301*	0.063	2	0.40	0.17
RM	29	11 (1)	1.338	1.217	0.246	0.250	0.145	0.208	-0.692*	-0.200	2	0.40	0.17
SDN	30	9	1.125	1.125	0.125	NA	0.063	NA	-1.000*	NA	1	0.00	0.00
Azores	119	21 (6)	2.414	2.159	0.169	0.181	0.267	0.330	0.366*	0.452*	2	0.19	0.08
CAR	30	15 (1)	1.708	1.307	0.175	0.177	0.235	0.308	0.257*	0.425*	1	0.00	0.00
CID	29	17 (2)	1.932	1.365	0.198	0.205	0.327	0.371	0.394*	0.447*	1	0.00	0.00
FOG	30	16 (1)	1.727	1.280	0.163	0.188	0.232	0.284	0.299*	0.340*	2	0.60	0.25
NAT	30	15	1.668	1.273	0.142	0.148	0.209	0.280	0.323*	0.472*	1	0.00	0.00
Canary Isl.	30	13	1.625	1.625	0.267	0.268	0.193	0.211	-0.385*	-0.268	1	0.00	0.00
IJU	30	13	1.498	1.216	0.267	0.268	0.193	0.211	-0.385*	-0.268	1	0.00	0.00
Cantabrian Cornice	169	26 (8)	2.745	2.461	0.133	0.248	0.369	0.473	0.639*	0.476*	3 (1)	0.42	0.18
BER	12	16 (1)	1.815	1.333	0.125	0.208	0.288	0.346	0.566*	0.398*	2	0.40	0.17
BAK	18	11	1.320	1.2	0.000	0.000	0.132	0.250	1.000*	1.000*	1	0.00	0.00
CUN	30	16 (1)	1.879	1.373	0.425	0.408	0.364	0.372	-0.169*	-0.096	1	0.00	0.00
EUM	30	8	1	1	0.000	NA	0.000	NA	NA	NA	1	0.00	0.00
LIE	30	9	1.125	1.125	0.125	NA	0.063	NA	-1.000*	NA	1	0.00	0.00
NUE	30	19 (1)	2.084	1.341	0.117	0.132	0.347	0.356	0.663*	0.630*	3 (1)	0.70	0.34
SEI	19	16 (1)	1.605	1.339	0.053	0.188	0.149	0.365	0.646*	0.486*	2	0.60	0.25
Total	443	37	1.520	1.239	0.153	0.177	0.528	0.545	0.186*	0.357*	4	0.52	0.0023

* $P < 0.05$; NA, not applicable

Plastid DNA sequence alignment included 82 sequences in total, with 235 base pairs (bp) in length and it included two variable positions. The total number of haplotypes found was four. Results for the diversity indices are shown in Table 3. At the population level, the mean diversity values for the ptDNA were *Hd* = 0.52 and π = 0.0023, with 12 populations showing null diversity values and the highest value in NUE (*Hd* = 0.7, π = 0.0034; Cantabrian Cornice). At the regional level, the most diverse regions were the Cantabrian Cornice and Andalusia (Table 3). Most of the populations showed only one haplotype.

3.2. Genetic structure and phylogeography

AMOVA analyses showed that, when all sampling units are included, the highest proportion of diversity lies in the interpopulation component (64.74%, *F* = 0.647), or in between regions when the regional component is considered (37.97%, *F* = 0.379) (Table 4). When we included only one individual per MLL per population, interpopulation or interregional variation decreased significantly, although they remained significantly high, and the highest proportion of variation resided within individuals.

Pairwise F_{ST} values with and without clonal individuals (excluding the monoclonal populations, with only one individual after clone correction) showed significant correlation (*r* = 0.945; *p* = 0.0001); lower and less paired significant differences were found when clones were excluded (Table S4). Significant differences were observed between most comparisons using all sampling units. At the intra-regional level, Andalusia and Cantabrian Cornice presented high F_{ST} values among several of their populations. Thus, in Andalusia, ALM and RM were very differentiated from each other and with respect to the population group CRM-SDN-PIN. In the Cantabrian Cornice, the populations or population groups EUM-SEI-BAK / NUE-CUN / BER / LIE showed high differentiation. Azorian populations showed low F_{ST} values between them. At inter-regional level most comparisons showed high differentiation, being lower between the Andalusian populations CRM-PIN-SDN and those of Azores, and the F_{ST} values were relatively low in the comparisons in which CUN was involved. The Canarian population (IJU) was very differentiated from the rest, except with PIN (Andalusia), the Azorian populations, and CUN.

Table 4. Hierarchical analysis of molecular variance (AMOVA).

Souce of variation	d.f.	Sum of squares	Percentage of variation	Phi	*P*-value
Microsatellites					
All sampling units					
Without grouping					
Within samples	443	547	29,23	0.708	<0.001
Between samples within population	426	742.883	6.03	0.171	<0.001
Between population	16	2296.697	64.74	0.647	<0.001
Total	885	3586.58	100		
Among regions (4 groups)					
Within samples	443	547	26.41	0.736	<0.001
Between samples within population	426	742.883	5.44	0.171	<0.001
Between population within region	13	949.498	30.18	0.486	<0.001
Between region	3	1347.199	37.97	0.379	<0.001
Total	885	3586.58	100		
MLLs					
Without grouping					
Within samples	130	224	42.6	0.574	<0.001
Between samples within population	113	371.188	19.31	0.312	<0.001
Between population within region	16	413.658	38.09	0.381	<0.001
Total	259	1008.846	100		
Among regions (4 groups)					
Within samples	130	224	40.44	0.596	<0.001
Between samples within population	113	371.188	18.33	0.312	<0.001
Between population within region	13	217.449	23.02	0.281	<0.001
Between region	3	196.208	18.21	0.182	<0.001
Total	259	1008.846	100		
cpDNA					
Without grouping					
Among populations	17	16.529	65.96		<0.001
Within populations	66	6.4	34.04		<0.001
Total	83	22.929	100		
Among regions (4 groups)					
Among groups	3	8.069	38.37		0.0123
Among populations within groups	14	7.314	30.44		<0.001
Within populations	66	6.200	31.19		<0.001
Total	83	21.583	100		
Among evolutionary units (2 groups)					
Among groups	1	7.873	50.53		0.0017
Among populations within groups	16	7.511	22.87		<0.001
Within populations	66	6.200	26.60		<0.001
Total	83	21.583	100		

The results of STRUCTURE with and without clones were very congruent with each other (Fig. S5); although without clones the structure is soon lost from $K = 8$, and the optimum number of selected clusters decreased from $K = 3$ and $K = 10$ (according to the ΔK and Puechmaille methods, respectively) with clones up to $K = 2$ and $K = 7$ without clones. When $K = 2$ (optimal K without clones; Fig. S5) one cluster was formed by the populations of Andalusia (less RM), Azores and Canaries, and another composed by the populations of the Cantabrian Cornice and RM (Andalusia); although NUE and CUN

(Cantabrian Cornice) had individuals better explained as belonging to the first cluster (Fig. 2). When K = 3 (optimal K with clones; Fig. S5), a third cluster related the populations of RM, PIN (11 individuals of 30), IJU, several individuals of SEI and BER, NUE (18 individuals of 30), and CUN. When considering the possibility of substructure (according to the Puechmaille method), for K = 10 the clusters obtained largely reflect the relationships obtained according to the pairwise F_{ST} values (i.e. ALM / CRM-PIN-SDN / RM / IJU-PIN / BER / EUM-SEI-BAK / LIE / NUE-CUN). The Azores populations appear differentiated from the rest, although with a great mix between them of other clusters. Thus, most of its individuals form two clusters (pink and green in Fig. 2), and many other individuals are better explained as belonging to clusters typical of other regions. In the Andalusian population PIN we see how there is a clear internal structure, that corresponds exactly with the existence of two intrapopulation nuclei (Fig. 2).

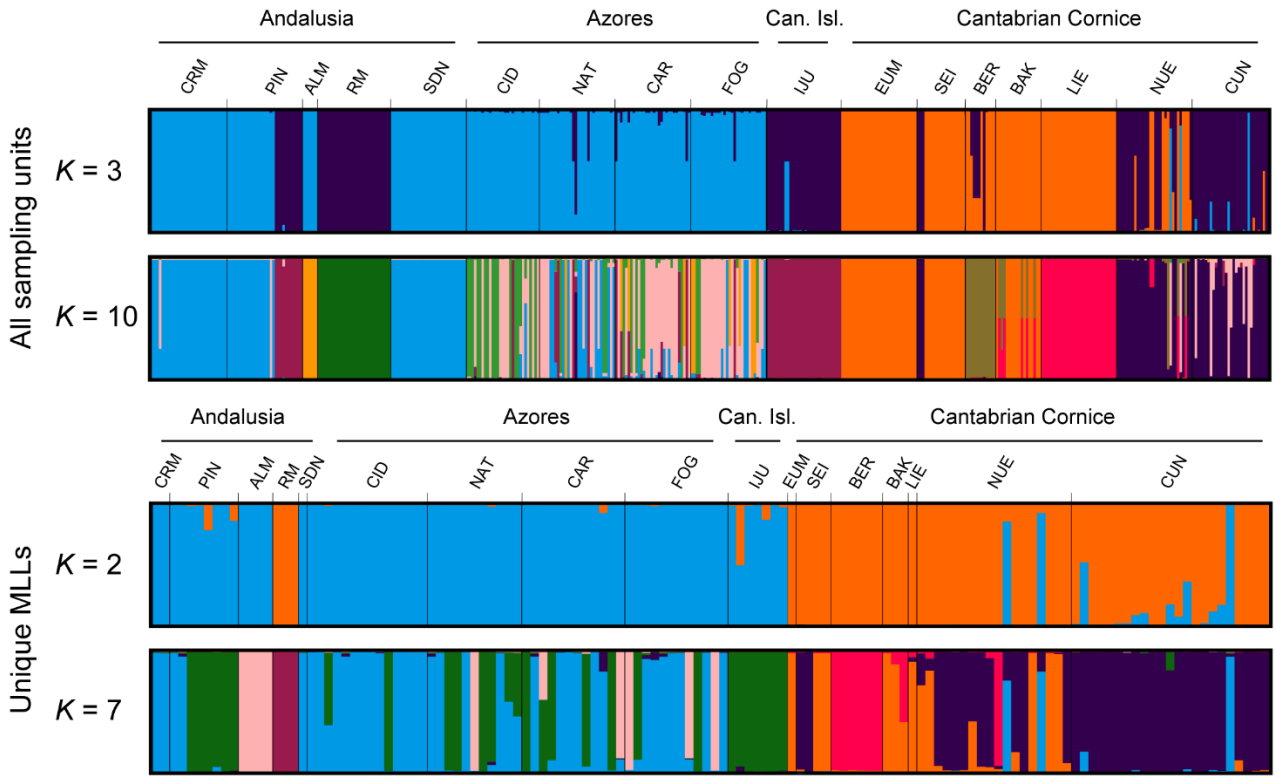


Figure 2. Estimated genetic structure based on microsatellite data using the Bayesian approach implemented in STRUCTURE. Histograms of individual assignment to clusters show the two most probable structuring, K = 3 and K = 11 for all sampling units, and K = 2 and K = 7 for unique MLLs.

The DAPC analysis resulted in relationships between the populations studied that reproduce greately the substructure obtained with STRUCTURE, and the relationships obtained with the pairwise F_{ST} values (Fig. 3). Briefly, on the one hand the well-

differentiated ALM, BER, LIE, and RM populations are observed, and on the other the groups formed by CUN-NUE, EUM-SEI-BAK, and IJU-PIN (individuals of an intrapopulational nucleus). The rest of Andalusian populations (CRM, SDN, rest of PIN) and the Azorians appear intermingled.

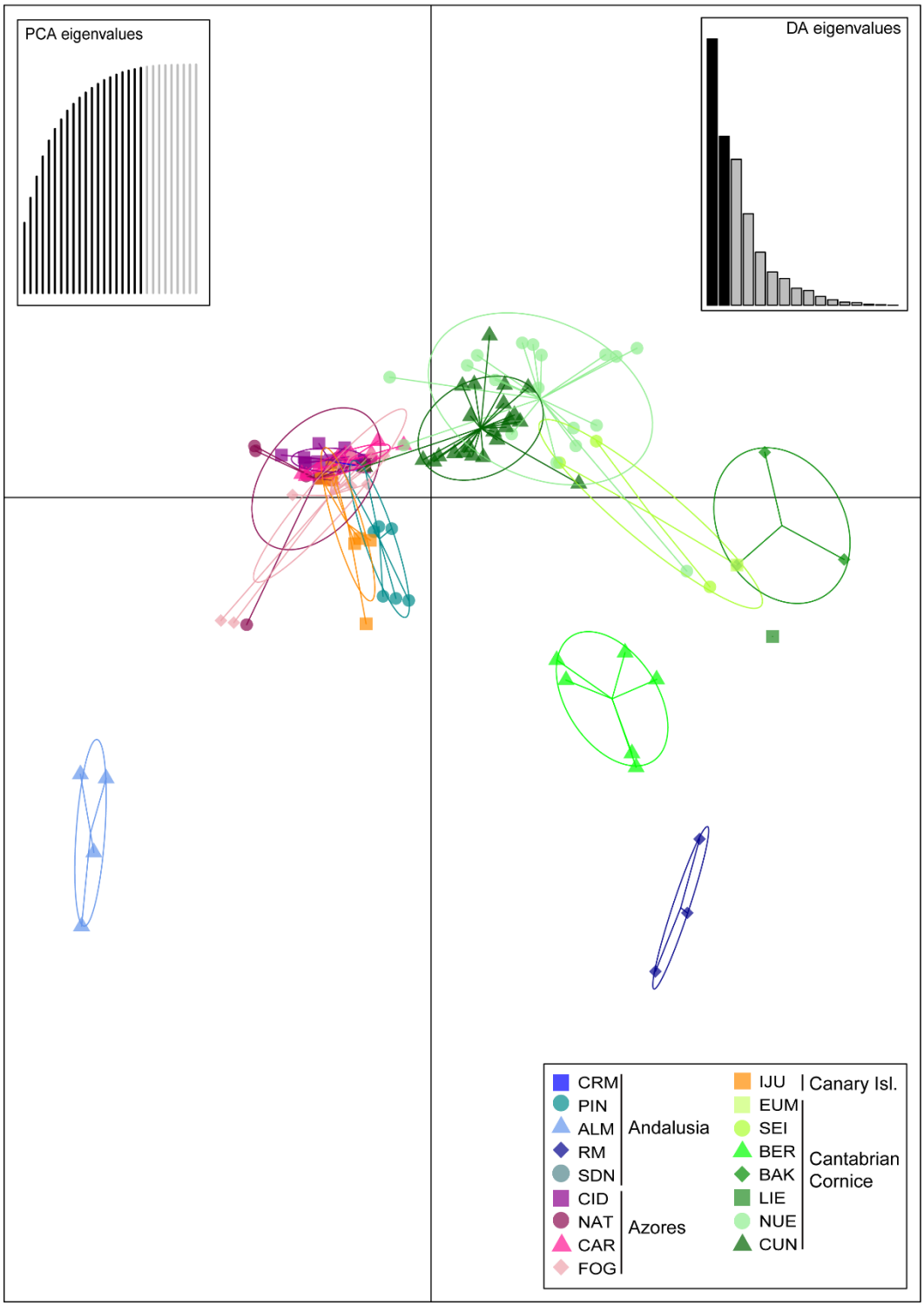


Figure 3: Result of the discriminant analysis of principal components (DAPC) using microsatellites.

The representation on a map of the ptDNA haplotype distributions suggests a geographical structuring of them (Figure 4a). Haplotypes H-I and H-II are the most frequent and widespread, showing a mainly different distribution. Haplotype H-I is dominant in Andalusia, Canary Islands and Azores (called Southern supra-regional group), but it also appears in northern Iberian (Cantabrian Cornice supra-regional group) where H-II is dominant. The latter also appears as the only one found in ALM (Andalusia) and together with H-I in FOG (Azores). The other two minority haplotypes are privates to Andalusia (H-III, in PIN and RM) and to the Cantabrian Cornice (H-IV, in NUE) (Figure 4a). PtDNA network (Fig. 4b) and phylogenetic analyses (Fig. 4c) suggested (with low posterior probabilities) haplotype clustering in two groups (H-I and H-III; H-II and H-IV), whose initial diversification began 19.5 Ma during the Burdigalian (Miocene) (Fig. 4c).

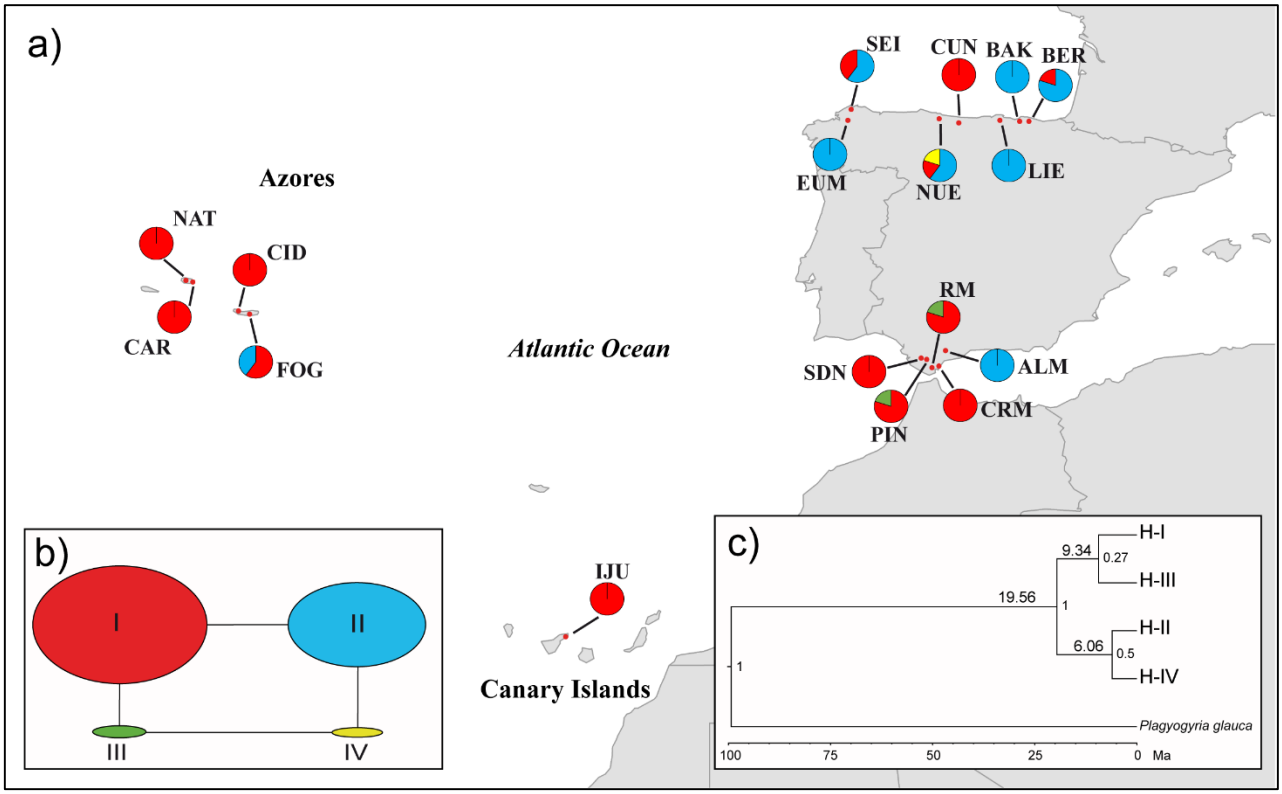


Figure 4: cpDNA information of the populations sampled. a) Distribution of the cpDNA haplotypes (pie charts), b) Inferred cpDNA network, following the statistical parsimony method, with TCS. Circle sizes are proportional to the haplotype frequencies. c) Dendrogram of the 4 haplotypes of the plastidial DNA of *C. macrocarpa* and the haplotype of *Plagiogyria glauca* used as an external group in the BEAST analysis. The divergence time (millions of years) of the nodes is shown and 95% of the highest posterior density (HPD) is indicated.

AMOVA analysis of ptDNA sequences showed that when considering 4 geographical regions (Andalusia, Azores, Canary Islands and Cantabrian Cornice), almost 39% (p = 0.0123) of variation was between regions (Table 4). Considering supra-regional grouping, (Southern and Cantabrian Cornice), a clear differentiation between them was evident (50.53%, p = 0.0017; Table 4).

3.3 Gene Flow

The results of BAYESASS indicated no current exchange of genes with the relative exception from CAR to NAT (migration rate [m] = 0.1357) and from CAR to FOG (m = 0.1335; Table S5). m estimates that did not exceed 0.110 (the upper value of the confidence interval when there is no information in the data) are considered that no current gene flow exist between populations.

The Mantel test, with and without clones, showed a lack of isolation by distance across the populations (r = 0.076, p = 0.209 with clones; r = 0.055, p = 0.304 without clones).

3.5 Species distribution modelling

For all models AUC values were high (minimum value of AUC = 0.990). The MAXENT current and LIG predictions showed regions of suitable habitats that coincided largely with the species' current distribution, with additional areas of it's distribution range in the European Atlantic coasts further north and the Mediterranean sea, where the species is currently absent (Fig. 5). According to LGM outputs, refugia were located in Macaronesia, the coast of Portugal and Galicia (northwest of Spain) (where *C. macrocarpa* is currently present) and the European Atlantic coast (at latitudes of the present north of France and South of Great Britain, where is currently absent). Paleodistribution modelling suggested no suitable habitats for *C. macrocarpa* in the northern coast of Spain (only in Galicia) were is currently present. The MAXENT future projections using the RCP8.5 scenario suggested a partial reduction of the suitable habitats at the coasts of Portugal, northern Iberian Peninsula, and Macaronesian islands along with an increase of the suitable habitats northwards of the European Atlantic coast.

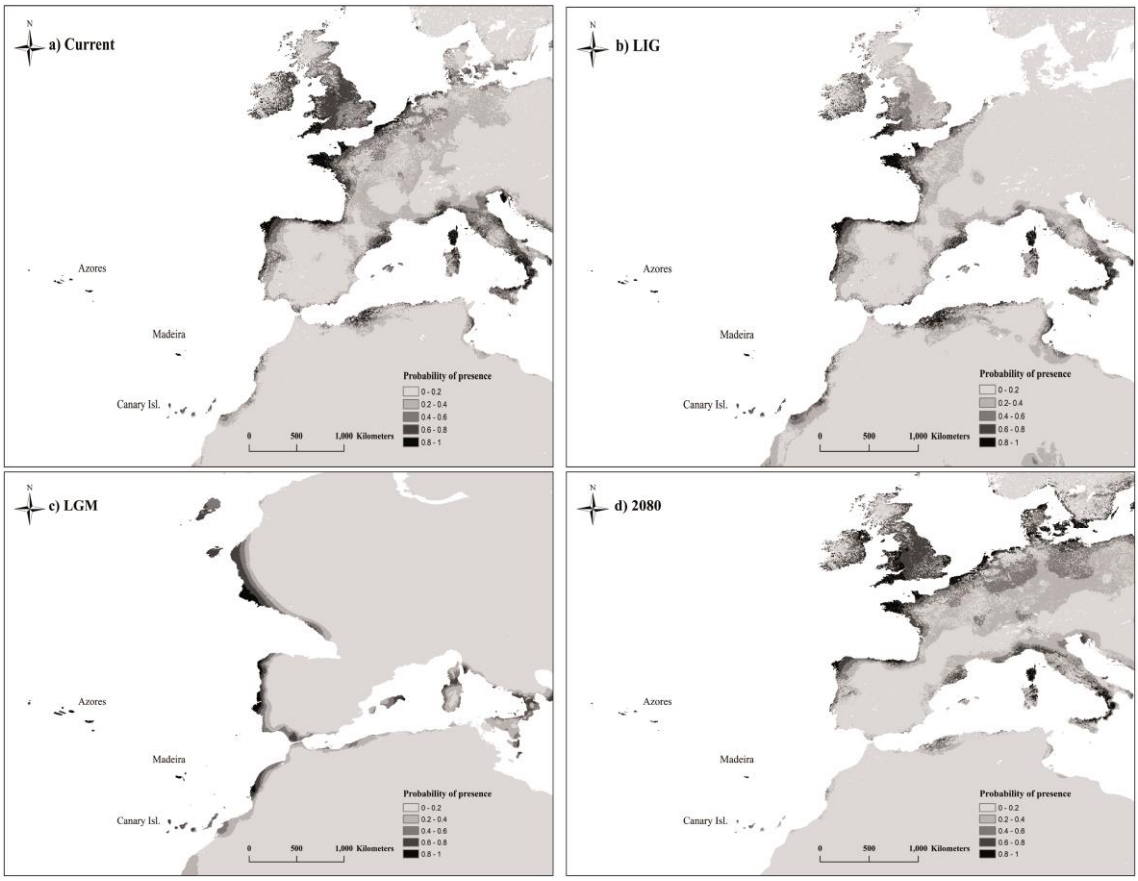


Figure 5: Potential distribution of *C. macrocarpa* drawn with MAXENT. Top left, at the present time; top right, at the Last Interglacial (LIG, ca. 120,000 years BP); bottom left, at the Last Glacial Maximum (LGM, ca. 21,000 years BP) using the Community Climate System Model (CCSM); bottom right, at 2.080 under RCP 8.5 conditions.

4. Discussion

Culcita macrocarpa has been considered one of the ferns of the herbaceous layer of European lauroid forests during the Tertiary and that, after the geological-climatic events of the Miocene and Pliocene, survived in Macaronesian and Iberian shelters (Benito Garzón & Sainz de Ollero; 2002). Few phylogeographic studies have been done on ferns (e.g. Trewick et al., 2002; Hunt et al., 2009; Jiménez et al., 2010; Bystriakova et al., 2014, Maccagni et al., 2017), and almost none focused on the history of ancestral taxa and the causes of their distribution models (e.g. Jiménez et al., 2010).

Our dating results of the cpDNA haplotypes confirm the Tertiary origin of *C. macrocarpa*, with the beginning of the diversification of gene lineages in the Burdigalian (19.5 Ma), about 4 million years after its divergence of *C. coniifolia* (23 Ma; Testo et al 2016). Recently, a third species was described, *C. remberi*, existing only in fossil material from the middle Miocene of the Clarkia fossil beds (16-15.4 Ma), in

northern Idaho (USA) (Pinson et al., 2018). Both *C. coniifolia* and *C. remberi* have an American distribution, with *C. coniifolia* growing in montane and submontane rainforest of Central and South America (Large and Braggins, 2004). Currently, *C. macrocarpa* is only present in Macaronesia and the Iberian Peninsula and no evidence is found about the existence in the American continent. Therefore, the origin of *C. macrocarpa* could be linked to a vicariance process, from the common ancestor with *C. coniifolia*, after the fragmentation of Laurasia (definitive at the end of the Paleogene).

The ability of ferns to disperse over long distances has caused their populations to be seen as exhibiting low genetic differentiation, with most of the variation residing at the intrapopulation level (Soltis & Soltis, 1989). Contrary to this, our results show a global phylogeography of *C. macrocarpa* characterized by the differentiation of populations in two main groups, which coincide with the geographical distribution of the species in the North of the Iberian Peninsula and in the Macaronesian archipelagos and southern Iberian Peninsula. In addition to this global model, we have also detected a strong interpopulation differentiation even within each of the two main groups, with the absence of recent gene flow. A similar phylogeographic model has been observed in *Vandenboschia speciosa* (Hymenophyllaceae), a Tertiary species with a distribution similar to that of *C. macrocarpa* (although more widespread northwards along the European Atlantic coast and towards central Europe) and whose populations are structured in two evolutionary units, one from the north (from the Cantabrian Cornice to the north and central Europe) and another from the south (Macaronesia, Andalusia and Italy) (Ben-Menni Schuler et al., in prep). In *V. speciosa*, the divergence between the gene lineages of the north and the south occurred coinciding with the Messinian salinity crisis (Ben-Menni Schuler et al., in prep). In the case of *C. macrocarpa*, the divergence of the two main haplotypes that characterize each population group began much earlier, during the Burdigalian, so the biogeographic history of this species has been marked by geological-climatic events since then.

Despite the influence of clonality on the levels of differentiation detected in *C. macrocarpa* with microsatellites, decreasing them (Tables 4, S4; Figs. 2, S5), both microsatellites and the cpDNA sequences coincided in the detection of the two main population groups (Figs. 2, 3; Table 4). When we integrate the data of genetic diversity, prevalence of clonal reproduction, and population differentiation, the picture that results is not consistent with the two population groups corresponding entirely to refuge

regions for *C. macrocarpa*. As expected for refuge areas (Provan and Bennett, 2008), the persistence of *C. macrocarpa* during climaticly unfavorable phases that affected Tertiary lauroid forests in Europe had to occur in the most diverse areas, with the highest prevalence of sexual reproduction and less population differentiation within these groups, from where it has subsequently dispersed generating less diverse populations, with a higher prevalence of vegetative reproduction and population differentiation. This model of low genetic diversity and high population differentiation has been demonstrated in various species of rock-dwelling ferns (e.g. Schneller and Holderegger, 1996; Vogel et al., 1999b; Suter et al., 2000). All populations in the Azores showed high relative diversity values (Table 3), low clonality (Fig. 1, Table 2), and it was the region with the least interpopulation differentiation and the only one within which recent gene flow was detected (Figs. 2, S4a, b, S5; Tables S4, S5). The same applies to the CUN and NUE populations from the Cantabrian Cornice. On the contrary, the Andalusian populations and the other populations of the Cantabrian Cornice showed low values of diversity (Table 3), strong impact of the clonality in their population configuration, with few clonal and very disproportionate lineages in terms of their intrapopulation dominance (Figs. 1, S4a, b; Table 2), and a clear differentiation of their populations (Figs. 2, 3, S5; Tables S4, S5), even among some very close ones (e.g. 2 km between CRM and RM) or between the two intrapopulation nuclei of PIN. These characteristics support the Azores and the CUN and NUE populations as shelters at least during the last glacial maximum, and that the Andalusian populations and the rest of the Cantabrian Cornice are the result of later dispersive events. These subsequent dispersive events are reflected in the distribution of the clonal lineages shared among populations, which have interregional structuring (Fig. 1). SDM results strongly support Azores as a refuge during the LGM and its presence during the LIG; however, they do not support the Cantabrian Cornise, but the most northwest end of the Iberian Peninsula (coast of Galicia) and the Portuguese coast. According to these results, *C. macrocarpa* would have recolonized the Cantabrica Cornice during interglacial periods and the Holocene. However, these results do not conform to the patterns of genetic diversity that we have obtained (little diversity in the Galician populations and much in the CUN and NUE populations of Asturias), which suggests that in the Cantabrian Cornice the populations persisted in small, climatically favorable pockets isolated from the most general climatic conditions. A similar pattern of diversity distribution and habitat suitability in the

Cantabrian Cornice was obtained for the also relict *V. speciosa* (Ben-Menni Schuler et al in prep).

The characteristics of populations with signs of dispersion are congruent with the occurrence of strong bottlenecks due to a recent founder event by one or few genotypes and subsequent expansion by vegetative reproduction, something known among herbaceous plants with clonal reproduction (e.g. Bauert et al., 1998; Paun et al., 2006) and among ferns in particular, where it is believed that most homospore ferns are capable of founding a population from a single spore via intragametophytic selfing (Peck et al., 1990 ; Schneller & Holderegger, 1996; Wolf et al., 2001). This form of colonization results in a totally homozygous sporophyte, generated from the same haploid gametophyte, and for a subsequent increase in genetic diversity and viability (especially for outcrossing diploids) sexual reproduction with additional immigrant genotypes is necessary (Pannell and Dorken, 2006). The founded populations EUM and BAK, with total homozygosity, and the rest of founded populations with total or almost total homozygosity after exclude locus *CM-AT19* (with fixed heterozygosity) could be examples of colonization from only one spore and establishment through intragametophytic selfing and vegetative reproduction. Our results agree with the culture experiments showing that the gametophytes of *C. macrocarpa* are initially male and later hermaphrodite (Stokey, 1930, Quintanilla et al., 2005), what seems to favor intragametophytic selfing (Klekowski and Lloyd, 1968). De Groot et al. (2012b) showed that intraspecific variation in mating system may be common, and that the genotypes with highest selfing capacity were those in isolated populations, supporting the idea that selection for selfing genotypes may occur during long-distance colonization. The high homozygosity observed in the founded populations of *C. macrocarpa* with respect to the source populations suggests a differential capacity for selfing, supporting the idea raised by De Groot et al. (2012b).

Deviations to the model of two main population groups are found in the analysis with STRUCTURE considering all sampling units (Fig. 2), which resulted in a third cluster linking two populations of the Cantabrian Cornice (CUN and NUE) with two Andalusian populations (RM and PIN) and with the Canarian population (IJU). This relationship was also reflected by the F_{ST} values (Table S4). In addition, although the two main haplotypes of cpDNA have a differentiated geographical distribution that characterizes the two population groups, the H-I haplotype occasionally appears in the

Cantabrian Cornice and the H-II does so in Andalusia and Azores. These exceptions suggest discrete long-distance dispersion events. The substructure detected in Andalusia and the different affinities (nuclear and chloroplastidial) of their populations with those of other regions suggest multiple colonizations of the region, from Azores and from the Cantabrian Cornice (or currently extinct populations that were closest to Andalusia). The RM population and one of the two intrapopulation nuclei of PIN are similar to the CUN and NUE populations with microsatellites, but show the typical cpDNA haplotype of the Macaronesian population group; However, the presence of this haplotype also in CUN and NUE suggests a long dispersion event towards Andalusia, where a local dispersion could subsequently take place from RM to PIN or vice versa. PIN and RM share the only private haplotype in Andalusia, supporting local dispersion. The presence of the haplotype H-II in ALM suggest another long-distance dispersal from the Cantabrian Cornice or from Azores (Fig. 4a), the latter being more likely given the nuclear relationship of ALM with Azores and other Andalusian populations when the uppermost hierarchical level of genetic structure is considered ($K = 2$, $K = 3$; Fig. 2).

In the Canary Islands only a population of *C. macrocarpa* is known, which due to intermediate levels of diversity and prevalence of clonality could represent a refuge population (Tables 2, 3), although it is also likely to be of recent colonization and that the relatively high levels of heterozygosity observed are due to accumulation of somatic mutations.

One limitation of our study, for a better interpretation of the connections between the regions of the Macaronesian group is the absence of sampling in Madeira. In the results obtained for *V. speciosa* Madeira was closely related to the south of the Iberian Peninsula, while Azores was closely related to the northern evolutionary unit (Ben-Menni Schuler et al., in prep). For *C. macrocarpa*, the populations of Azores are clearly related to the south of the Iberian Peninsula and the Canary Islands. The absence of Madeira samples prevents evaluating the possible role of this archipelago in facilitating the connection of Azores with the Canary Islands and the south of the Iberian Peninsula.

Conservation and future perspectives

The identificacion of two supraregional groups (Cantabrian cornice/South) could led to design a management plan to improve the conservation of the species. This fern is

specially threatened in Andalusia, where private haplotypes and alleles are present, being important the protection of this area. In relation to future predictions, SDM results in 2080 show a very accused loss in habitat suitability for *C. macrocarpa* in Andalusia and in the Canary Islands being threatened the southern group. In the northern group, Galicia also shows a decrease in habitat suitability. On the other hand, a northward migration is proposed for the species. All this makes necessary the protection of the southern populations to try to prevent the loss of genetic diversity.

References

Agapow, P.M., Burt, A. 2001. Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes* 1: 101–102.

Anonymous. 1979. Convention of the conservation of European wildlife and natural habitats. *European Treaty Series* 104: 1–27.

Anonymous. 1992. Council Directive 92/43/EEC of 21 May 1992 on the conservation of habitats and wild fauna and flora. *Official Journal of the European Communities L* 206: 7–50.

Arnaud-Haond, S., Duarte, C.M., Alberto, F., Serrao, E.A., 2007. Standardizing methods to address clonality in population studies. *Molecular ecology* 16*:* 5115-5139.

Arseneau, J.R., Steeves, R., Laflamme, M. 2017. Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues. *Molecular ecology resources* 17: 686-693.

Balloux, F., Lehmann, L., de Meeûs, T. 2003. The population genetics of clonal and partially clonal diploids. *Genetics* 164: 1635-1644.

Bañares, A., Blanca, G., Güemes, J., Moreno, J. C., Ortiz, S. 2008. *Lista roja 2008 de la flora vascular española.* Gen. de Medio Natural y Política Forestal (Min. de Medio Ambiente, y Medio Rural y Marino)-SEBICOP, Madrid.

Barrón, E. 2003. Evolución de las floras terciarias en la Península Ibérica. *Monografías del Jardín Botánico de Córdoba* 11: 63–74.

Barrón, E., Peyrot, D. 2006. La vegetación forestal en el Terciario. In: Carrión J, Fernández S, Fuentes N, eds. *Paleoambientes y cambio climático*. Murcia, Spain: Fundación Séneca/Agencia de Ciencia y Tecnología de la Región de Murcia, 56–77.

Benito-Garzón, M., Sainz de Ollero, H. 2002. *Potencialidad del elemento paleotropical lauroide en la Península Ibérica.* MSc thesis, Universidad Autónoma de Madrid, Madrid, Spain.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS computational biology* 10: e1003537.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS computational biology* 15: e1006650.

Bystriakova, N., Ansell, S.W., Russell, S.J., Grundmann, M., Vogel, J.C., Schneider, H. 2014. Present, past and future of the European rock fern *Asplenium fontanum*: combining distribution modelling and population genetics to study the effect of climate change on geographic range and genetic diversity. *Annals of Botany* 113: 453–465.

Carrillo-Angeles, I.G., Golubov, J., Milligan, B.G., Mandujano, M.C. (2011). Spatial distribution pattern of a clonal species: effects of differential production of clonal and sexual offspring. *Evolutionary ecology* 25: 1357-1383.

Chen, C., Qi, Z.C., Xu, X.H., Comes, H.P., Koch, M.A., Jin, X.J., Fu, C.X., Qiu, Y.X. 2014. Understanding the formation of Mediterranean–African–Asian disjunctions: evidence for Miocene climate-driven vicariance and recent long-distance dispersal in the Tertiary relict *Smilax aspera* (Smilacaceae). *New Phytologist* 204: 243–255.

Clement, M., Posada, D., Crandall, K.A. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9: 1657–1659.

Comes, H.P., Kadereit, J.W. 1998. The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in plant science* 3: 432-438.

Cruzan, M.B., Templeton, A.R. 2000. Paleoecology and coalescence: phylogeographic analysis of hypotheses from the fossil record. *Trends in Ecology & Evolution* 15: 491-496.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature methods* 9: 772.

De Meeûs, T., Lehmann, L., Balloux, F. 2006. Molecular epidemiology of clonal diploids: a quick overview and a short DIY (do it yourself) notice. *Infection, Genetics and Evolution* 6: 163-170.

Dorken, M.E., Eckert, C.G. 2001. Severely reduced sexual reproduction in northern populations of a clonal plant, Decodonverticillatus (Lythraceae). *Journal of Ecology* 89: 339-350.

Douhovnikoff, V., Leventhal, M. 2016. The use of Hardy–Weinberg Equilibrium in clonal plant systems. *Ecology and evolution* 6: 1173-1180.

Drummond, A.J., Ho, S.Y., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. *PLoS biology* 4: e88.

Evanno, G., Regnaut, S. Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular ecology* 14:2611-2620.

Excoffier, L. Lischer, H.E., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources* 10: 564-567.

Excoffier, L., Smouse, P.E. Quattro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.

Falush, D., Stephens, M. Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567-1587.

García-Verdugo, C., Calleja, J.A., Vargas, P., Silva, L., Moreira, O., Pulido, F. 2013. Polyploidy and microsatellite variation in the relict tree *Prunus lusitanica* L.:

how effective are refugia in preserving genotypic diversity of clonal taxa? *Molecular Ecology* 22: 1546–1557.

Grünwald, N.J., Goodwin, S.B., Milgroom, M.G., Fry, W.E. 2003. Analysis of genotypic diversity data for populations of microorganisms. *Phytopathology* 93*:* 738–746.

Gomez, A., Wright, P.J., Lunt, D.H., Cancino, J.M., Carvalho, G.R., Hughes, R.N. 2006. Mating trials validate the use of DNA barcoding to reveal cryptic speciation of a marine bryozoan taxon. *Proceedings of the Royal Society B: Biological Sciences* 274: 199-207.

Goudet, J. 2005. Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes* 5: 184-186.

De Groot, G.A., During, H.J., Ansell, S.W., Schneider, H., Bremer, P., Wubs, E.J., Maas, J.W., Korpelainen, H. Erkens, R.H. 2012b. Diverse spore rains and limited local exchange shape fern genetic diversity in a recently created habitat colonized by long-distance dispersal. *Annals of Botany* 109: 965-978.

de Groot, G. A., Verduyn, B., Wubs, E. J., Erkens, R. H., During, H. J. 2012a. Inter-and intraspecific variation in fern mating systems after long-distance colonization: the importance of selfing. *BMC Plant Biology* 12: 3.

Hall, T.A. 1999. January. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic acids symposium series* (Vol. 41, No. 41, pp. 95-98). [London]: Information Retrieval Ltd., c1979-c2000.

Hampe, A., Petit, R. J. 2005. Conserving biodiversity under climate change: the rear edge matters. *Ecology letters* 8: 461-467.

Hengl, T., de Jesus, J.M., MacMillan, R.A., Batjes, N.H., Heuvelink, G.B., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J.G., Walsh, M.G., Gonzalez, M.R., 2014. SoilGrids1km—global soil information based on automated mapping. *PloS one* 9: e105992.

Hewitt, G.M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* 58: 247–276.

Hewitt, G.M. 1999. Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* 68: 87–112.

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: A Journal of the Royal Meteorological Society* 25: 1965–1978.

Hunt, H.V., Ansell, S.W., Russell, S.J., Schneider, H., Vogel, J.C. 2009. Genetic diversity and phylogeography in two diploid ferns, *Asplenium fontanum* subsp. *fontanum* and *A. petrarchae* subsp. *bivalens*, in the western Mediterranean. *Molecular Ecology* 18: 4940–4954

Hurlbert, S.H. 1971. The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 52: 577-586.

Jiménez, A., Holderegger, R., Csencsics, D., Quintanilla, L.G. 2010. Microsatellites reveal substantial among-population genetic differentiation and strong inbreeding in the relict fern *Dryopteris aemula*. *Annals of Botany* 106: 149–155.

Jombart, T., Devillard, S. Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics* 11: 94.

Kamvar, Z.N., Tabima, J.F. Grünwald, N.J., 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2: e281.

Klekowski Jr, E.J., Lloyd, R.M. 1968. Reproductive biology of the Pteridophyta: 1. General considerations and a study of *Onoclea sensibilis* L. *Botanical Journal of the Linnean Society* 60: 315-324.

Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., Mayrose, I., 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular ecology resources* 15: 1179-1191.

Li, Y.L. Liu, J.X., 2018. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Molecular Ecology Resources* 18: 176-177.

J.A Ludwig, J.F Reynolds. 1988. Statistical Ecology, Wiley, New York.

Maccagni, A., Parisod C, Grant J.R. 2017. Phylogeography of the moonwort fern *Botrychium lunaria* (Ophioglossaceae) based on chloroplast DNA in the Central-European Mountain System. *Alpine Botany* 127: 185–196.

Mai, D.H. 1989. Development and regional differentiation of the European vegetation during the Tertiary. *Plant Systematics and Evolution* 162: 79–91.

Martínez, B., Viejo, R.M., Carreño, F., Aranda, S.C. 2012. Habitat distribution models for intertidal seaweeds: responses to climatic and non-climatic drivers. *Journal of Biogeography* 39: 1877–1890.

Médail, F., Diadema, K. 2009. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of biogeography* 36: 1333-1345.

Meirmans, P.G. Van Tienderen, P.H. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4: 792-794.

Nei, M. 1987. *Molecular evolutionary genetics*, vol 512. Columbia University Press, New York

Petit, R. J., Aguinagalde, I., de Beaulieu, J. L., Bittkau, C., Brewer, S., Cheddadi, R., ... & Mohanty, A. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. *science* 300: 1563-1565.

Petit, R.J., Hampe, A., Cheddadi, R. 2005. Climate change and tree phylogeography in the Mediterranean. *Taxon* 54: 877–885.

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

Pichi-Sermolli, R.E.G. 1979. A survey of the pteridological flora of the Mediterranean region. *Webbia* 34: 175–242.

Pichi-Sermolli, R.E.G. 1991. On the taxonomy and nomenclature of some species from tropical Africa of the genus *Triplophyllum* Holttum (Dryopteridaceae). *Webbia* 45: 117–135.

Pielou, E.C. 1975. *Ecological Diversity*. Wiley Interscience, New York, NY, U.S.A.

Provan, J., Bennett, K.D. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology & Evolution* 23: 564–571.

Puechmaille, S.J. 2016. The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular Ecology Resources* 16: 608-627.

Quintanilla, L.G. Amigo, J. 2001. *Culcita macrocarpa* C. Presl (Dicksoniaceae, Pteridophyta) en la cuenca del Xubia (NO de la Peninsula Iberica). *Nova Acta Científica Compostelana (Biol.)* 11: 267-268.

Quintanilla, L.G., Pajarón, S., Pangua, E., Amigo, J. 2000. Effect of temperature on germination in northernmost populations of Culcita macrocarpa and Woodwardia radicans. *Plant Biology* 2: 612-617.

Quintanilla, L.G., Pajarón, S., Pangua, E., Amigo, J. 2007. Allozyme variation in the sympatric ferns Culcita macrocarpa and Woodwardia radicans at the northern extreme of their ranges. *Plant Systematics and Evolution* 263: 135-144.

Quintanilla, L.G., Pangua, E., Amigo, J., Pajarón, S. 2005. Comparative study of the sympatric ferns Culcita macrocarpa and Woodwardia radicans: sexual phenotype. *Flora-Morphology, Distribution, Functional Ecology of Plants* 200: 187-194.

Rodríguez-Sánchez, F., Guzmán, B., Valido, A., Vargas, P., Arroyo, J. 2009. Late Neogene history of the laurel tree (*Laurus* L., Lauraceae) based on phylogeographical analyses of Mediterranean and Macaronesian populations. *Journal of Biogeography* 36: 1270–1281.

Rodríguez Guitián, M. A., Ferreiro da Costa, J., Ramil Rego, P., Lijó Pose, G. 2011. Caracterización ambiental, demografía y amenazas para su conservación de la población lucense de *Culcita macrocarpa* C. Presl.(NW ibérico). *Recursos Rurais* 7: 15-25

Romero, M.I., Rodríguez Guitián, M.A. Rubinos, M. 2004: Adiciones al catálogo pteridológico gallego. *Botanica Complutensis* 28: 51-55.

Shepherd, L.D., Perrie, L.R., Brownsey, P.J. 2008. Low-copy nuclear DNA sequences reveal a predominance of allopolyploids in a New Zealand Asplenium fern complex. *Molecular Phylogenetics and Evolution* 49: 240-248.

Silvertown, J. 2008. The evolutionary maintenance of sexual reproduction: evidence from the ecological distribution of asexual reproduction in clonal plants. *International Journal of Plant Sciences* 169: 157-168.

Simpson, E.H. 1949. Measurement of diversity. *Nature* 163: 688.

Soltis, D.E., Soltis, P.S. 1989. Polyploidy, breeding systems, and genetic differentiation in homosporous pteridophytes. In: Soltis, D. E. and Soltis, P. S. (Eds.). *Isozymes in plant biology*. Oregan: Dioscorides Press.v. 4: 241-258.

Soñora, F.X. 1992. Notas pteridológicas de Galicia. IV.*Acta Botanica Malacitana* 17: 282-283.

Stenberg, P., Lundmark, M. Saura, A., 2003. MLGsim: a program for detecting clones using a simulation approach. *Molecular Ecology Notes* 3: 329-331.

Stewart, J.R., Lister, A.M. 2001. Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology & Evolution* 16: 608–613.

Suter, M., Schneller, J.J., Vogel, J.C. 2000. Investigations into the genetic variation, population structure, and breeding systems of the fern Asplenium trichomanes subsp. quadrivalens. *International Journal of Plant Sciences* 161: 233-244.

Stokey, A.G. 1930. Prothallia of the Cyatheaceae. *Botanical Gazette* 90: 1-45.

Taberlet. P., Fumagalli. L., Wust-Saucy. A.G, Cossons. J.F. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* 7: 453– 464.

Templeton, A.R., Crandall, K.A., Sing, C.F. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132: 619– 633.

Testo, W. Sundue, M., 2016. A 4000-species dataset provides new insight into the evolution of ferns. *Molecular Phylogenetics and Evolution* 105: 200-211.

Trewick, S.A., Morgan-Richards, M., Russell, S.J., Henderson, S., Rumsey, F.J., Pintér, I., Barret, J.A., Gibby, M., Vogel, J.C. 2002. Polyploidy, phylogeography and Pleistocene refugia of the rockfern *Asplenium ceterach*: evidence from chloroplast DNA. *Molecular Ecology* 11: 2003–2012.

Van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque J-F et al. 2011. The representative concentration pathways: an overview. *Climatic Change* 109: 5.

Vogel, J.C., Rumsey, F.J., Schneller, J.J., Barrett, J.A., Gibby, M. 1999. Where are the glacial refugia in Europe? Evidence from pteridophytes. *Biological Journal of the Linnean Society* 66: 23–37.

Wang, L., Wu, Z. Q., Bystriakova, N., Ansell, S. W., Xiang, Q. P., Heinrichs, J., ... & Zhang, X.C. 2011. Phylogeography of the Sino-Himalayan fern Lepisorus clathratus on "the roof of the world". *PloS one* 6: e25896.

Weir, B. S., Cockerham, C. C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.

Wilson, G.A. Rannala, B., 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163: 1177-1191.

Wolf, P.G., Der, J.P., Duffy, A.M., Davidson, J.B., Grusz, A.L., Pryer, K.M. 2011. The evolution of chloroplast genes and genomes in ferns. *Plant molecular biology* 76: 251-261.

Supporting information

Table S1. Primer pairs used for PCR amplification of the 13 cpDNA regions tested in *C. macrocarpa*.

Primer	**Sequences 5'-3'**	**Region**
rpL16-F-Culcit	ATGCCTAGTGTGCGACCCGTT	*rpl16* intron
rpL16-R-Culcit	TCCTCTATGTTGCTTACGGAAT	*rpl16* intron
CM-rpS16-F	AAGCGACATGGTGGGAAGCAA	*rps16* intron
CM-rpS16-R	CGGGACCGAGCATCAATTGCTA	*rps16* intron
CM-ycf3F	GCTTCTACATATTATAGGGATG	*ycf3* introns
CM-ycf3R	TTGAATGGCCTGTTCCCC	*ycf3* introns
trnL(UAG)	CTGCTTCCTAAGAGCAGCGT	*rpl32-trnL*
rpL32-F	CAGTTCCAAAGAAACGTACCTC	*rpl32-trnL*
CM-rps4-trnL(CAA)-Fw	CTTCTCCTGGATTGGATGCT	*rps4-trnL* spacer
CM-rpS4-trnL(CAA)-Rv	CGCAGCGTCTACCGTTCCG	*rps4-trnL* spacer
CM-rps16-matK-Fw	CATATTGAAACGGGCAAAGG	*rps16-matK*
CM-rps16-matK-Rev	CTTTTCCGTAACGTCCCAAA	*rps16-matK*
CM-trnD-psbM-F	ACCGATTGAACTACAATCCC	*trnD*(GUC)-*psbM*
CM-trnD-psbM-R	GGCTACGAACGCAAGAATATTGACTTCCGT	*trnD*(GUC)-*psbM*
CM-trnD(GUC)-R	GGGATTGTAGTTCAATCGGT	*trnD-trnY-trnE-rpoB*
CM-rpoB	CGGTTAGATCCCTCAAATTG	*trnD-trnY-trnE-rpoB*
CM-trnH(GUG)2	TGGATCCACAATCCATTGC	*trnH-psbA*
CM-psbA2	CGTAACGCTCATAACTTCCCTCT	*trnH-psbA*
f	ACTTGAACTGGTGACACGAG	*trnL-trnL-trnF*; Taberlet et al., 1991
e	GGTTCAAGTCCCTCTATCCC	*trnL-trnF*; Taberlet et al., 1991
CM-trnL-Fw	CGGAACGGTAGACGCTGCG	*trnL-trnL-trnF*
trnS(GCU)	AGAGAGGGATTCGAACCCTCGGT	*trnS-trnG-trnG*; Shaw et al., 2005
3'trnG(UUC)	GTAGCGGGAATCGAACCCGCATC	*trnS-trnG-trnG*; Shaw et al., 2005
5'trnG2S	TTTTACCACTAAACCATACCCGC	*trnS-trnG*; Shaw et al., 2005
5'trnG2G	GCGGGTATGGTTTAGTGGTAAAA	trnG-trnG; Shaw et al., 2005

Table S2. Percent contribution and permutation importance (MaxEnt) of selected model. Variables in bold were selected for the final model.

Variable	MaxEnt Percent contribution	MaxEnt Permutation importance
Mean Diurnal Range	48	11
Min. Temperature of Coldest Month	30.9	80.1
Type of Soil	10.6	1.4
Precipitation of Warmest Quarter	8.4	5
Precipitation of Coldest Quarter	1	0.1
Max. Temperature of Warmest Month	0.7	2
Precipitation Seasonality	0.3	0.4
Precipitation of Wettest Month	0.1	0.2
Precipitation of Driest Month	0	0

Table S3. Values for F_{IS} per population and per locus with and without locus *CM-AT19*.

	With *CM-AT19*									Without *CM-AT19*							
Population	*CM1A*	*CM35*	*CM21b*	*CM-AT30*	*CM-AT45m1*	*CM-AT9*	*CM-AT2*	*CM-AT19*	Multi-locus	*CM1A*	*CM35*	*CM21b*	*CM-AT30*	*CM-AT45m1*	*CM-AT9*	*CM-AT2*	Multi-locus
CRM	---	---	---	---	---	---	0.000	-1.000	-0.938*	---	---	---	---	---	---	0.000	0.000
PIN	---	---	---	0.776	---	0.861	0.691	-1.000	0.301*	---	---	---	0.776	---	0.861	0.691	0.779*
ALM	---	---	---	---	---	---	0.750	-1.000	0.000	---	---	---	---	---	---	0.750	0.750*
RM	---	---	-0.867	---	---	---	0.788	-1.000	-0.692*	---	---	-0.867	---	---	---	0.788	-0.460*
SDN	---	---	---	---	---	---	---	-1.000	-1.000*	---	---	---	---	---	---	---	---
CID	---	1.000	0.154	0.729	---	1.000	0.853	-1.000	0.394*	---	1.000	0.154	0.729	---	1.000	0.853	0.723*
NAT	---	---	1.000	---	1.000	1.000	0.779	-1.000	0.323*	---	---	1.000	---	1.000	1.000	0.779	0.887*
CAR	---	---	0.437	---	0.846	1.000	0.557	-1.000	0.257*	---	---	0.437	---	0.846	1.000	0.557	0.711*
FOG	---	---	-0.036	0.728	1.000	1.000	0.721	-1.000	0.299*	---	---	-0.036	0.728	1.000	1.000	0.721	0.779*
EUM	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
SEI	1.000	---	0.854	0.786	-0.029	0.000	---	-0.059	0.646*	1.000	---	0.854	0.786	-0.029	0.000	---	0.747*
IJU	---	---	---	0.120	---	-0.383	---	-1.000	-0.385*	---	---	---	0.120	---	-0.383	---	-0.090
BER	0.656	---	---	1.000	0.560	---	0.814	-0.375	0.566*	0.656	---	---	1.000	0.560	---	0.814	0.778*
BAK	1.000	---	---	1.000	---	---	1.000	---	1.000*	1.000	---	---	1.000	---	---	1.000	1.000*
LIE	---	---	---	---	---	---	---	-1.000	-1.000*	---	---	---	---	---	---	---	---
NUE	1.000	0.439	0.000	0.630	0.870	0.867	1.000	-0.137	0.663*	1.000	0.439	0.000	0.630	0.870	0.867	1.000	0.737*
CUN	0.000	-0.674	0.164	0.100	0.138	0.065	---	-0.758	-0.169*	0.000	-0.674	0.164	0.100	0.138	0.065	---	-0.049
Overall	0.898	0.125	0.085	0.634	0.609	0.636	0.748	-0.894	0.186	0.898	0.125	0.085	0.634	0.609	0.636	0.748	0.578

*, $P < 0.05$

Table S4. Pairwise population *F*ST for microsatellites. *F*ST values with all sampled individuals are represented below de diagonal, and *F*ST values with a single representing MLL per population above the diagonal. Values in bold were significant at the 5% nominal level after sequential Bonferroni correction.

	CRM	PIN	ALM	RM	SDN	CID	NAT	CAR	FOG	EUM	SEI	IJU	BER	BAK	LIE	NUE	CUN
CRM	--	0.231	0.611	0.700	0	0.011	-0.056	-0.016	-0.076	0	0.589	0.519	0.591	0.767	0	0.469	0.259
PIN	**0.192**	--	0.525	0.602	0	0.221	**0.196**	**0.225**	0.197	0	0.528	**0.274**	**0.517**	0.694	0	**0.471**	**0.255**
ALM	**0.800**	**0.531**	--	0.774	0	**0.437**	0.407	**0.410**	0.407	0	0.638	0.621	0.650	0.781	0	**0.550**	**0.437**
RM	**0.829**	**0.678**	**0.820**	--	0	**0.500**	**0.611**	**0.593**	**0.590**	0	0.551	**0.678**	0.629	0.715	0	**0.566**	**0.466**
SDN	0.000	**0.199**	**0.809**	**0.833**	--	0	0	0	0	0	0	0	0	0	0	0	0
CID	**0.342**	**0.204**	**0.476**	**0.605**	**0.350**	--	0.110	0.101	0.081	0	**0.490**	**0.344**	**0.493**	0.615	0	**0.393**	**0.230**
NAT	**0.190**	0.096	**0.512**	**0.704**	**0.202**	**0.151**	--	-0.017	0.018	0	0.565	**0.301**	**0.572**	0.698	0	**0.446**	**0.237**
CAR	**0.323**	**0.170**	**0.481**	**0.679**	**0.335**	**0.1221**	0.028	--	0.023	0	**0.547**	**0.303**	**0.536**	0.667	0	**0.399**	**0.221**
FOG	**0.259**	**0.129**	**0.472**	**0.677**	**0.272**	**0.123**	0.026	0.023	--	0	0.547	**0.326**	**0.570**	0.692	0	**0.435**	**0.197**
EUM	**0.964**	**0.865**	**0.972**	**0.876**	**0.967**	**0.805**	**0.879**	**0.857**	**0.864**	--	0	0	0	0	0	0	0
SEI	**0.888**	**0.755**	**0.818**	**0.745**	**0.891**	**0.681**	**0.774**	**0.744**	**0.753**	0.137	--	**0.593**	0.434	0.227	0	0.248	**0.268**
IJU	**0.707**	**0.448**	**0.646**	**0.732**	**0.714**	**0.417**	**0.448**	**0.392**	**0.431**	**0.882**	**0.764**	--	**0.612**	0.756	0	**0.497**	**0.302**
BER	**0.819**	**0.638**	**0.696**	**0.739**	**0.823**	**0.557**	**0.675**	**0.641**	**0.657**	**0.756**	**0.530**	**0.681**	--	0.367	0	**0.408**	**0.481**
BAK	**0.903**	**0.777**	**0.842**	**0.778**	**0.906**	**0.704**	**0.795**	**0.767**	**0.776**	0.349	0.126	**0.795**	**0.523**	--	0	0.321	**0.505**
LIE	**0.920**	**0.810**	**0.895**	**0.821**	**0.923**	**0.746**	**0.817**	**0.788**	**0.802**	**0.833**	**0.587**	**0.848**	**0.710**	**0.616**	--	0	0
NUE	**0.684**	**0.540**	**0.570**	**0.616**	**0.689**	**0.418**	**0.531**	**0.473**	**0.508**	**0.573**	**0.391**	**0.526**	**0.452**	**0.404**	**0.530**	--	**0.214**
CUN	**0.496**	**0.305**	**0.439**	**0.529**	**0.504**	**0.226**	**0.281**	**0.221**	**0.229**	**0.694**	**0.523**	**0.352**	**0.518**	**0.571**	**0.628**	**0.241**	--

Table S5. Mean recent migration rates (m) among the studied populations, estimated from eight microsatelliteloci using the BAYESASS program. Values on the diagonal (underlined) indicate the proportion of individuals in each generation that are not migrants. Values in bold are the *m* rates that are informative.

	From																
To	CRM	PIN	ALM	RM	SDN	CID	NAT	CAR	FOG	EUM	SEI	IJU	BER	BAK	LIE	NUE	CUN
CRM	0.684	0.017	0.018	0.016	0.016	0.018	0.02	0.048	0.017	0.017	0.018	0.017	0.018	0.016	0.019	0.017	0.016
	(0.016)	(0.017)	(0.017)	(0.016)	(0.016)	(0.016)	(0.019)	(0.027)	(0.016)	(0.015)	(0.018)	(0.017)	(0.017)	(0.015)	(0.017)	(0.016)	(0.015)
PIN	0.013	0.680	0.013	0.014	0.013	0.013	0.015	0.038	0.013	0.013	0.013	0.092	0.012	0.013	0.012	0.014	0.013
	(0.014)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)	(0.014)	(0.022)	(0.013)	(0.012)	(0.012)	(0.031)	(0.012)	(0.012)	(0.011)	(0.013)	(0.013)
ALM	0.015	0.015	0.745	0.015	0.016	0.016	0.017	0.014	0.015	0.015	0.016	0.016	0.016	0.015	0.015	0.015	0.016
	(0.015)	(0.014)	(0.030)	(0.014)	(0.015)	(0.016)	(0.016)	(0.014)	(0.013)	(0.015)	(0.016)	(0.017)	(0.015)	(0.013)	(0.014)	(0.015)	(0.015)
RM	0.0154	0.015	0.014	0.756	0.015	0.014	0.015	0.015	0.014	0.014	0.016	0.015	0.015	0.015	0.015	0.015	0.015
	(0.015)	(0.014)	(0.014)	(0.030)	(0.015)	(0.012)	(0.013)	(0.015)	(0.015)	(0.014)	(0.015)	(0.014)	(0.015)	(0.015)	(0.014)	(0.015)	(0.013)
SDN	0.018	0.017	0.018	0.018	0.685	0.017	0.021	0.034	0.018	0.019	0.018	0.018	0.018	0.018	0.018	0.019	0.018
	(0.017)	(0.016)	(0.016)	(0.017)	(0.017)	(0.016)	(0.02)	(0.024)	(0.017)	(0.017)	(0.018)	(0.018)	(0.016)	(0.016)	(0.016)	(0.018)	(0.017)
CID	0.009	0.009	0.010	0.010	0.01	0.749	0.013	0.094	0.009	0.01	0.01	0.01	0.009	0.011	0.011	0.01	0.01
	(0.008)	(0.009)	(0.011)	(0.010)	(0.01)	(0.029)	(0.014)	(0.030)	(0.009)	(0.009)	(0.01)	(0.01)	(0.009)	(0.01)	(0.01)	(0.012)	(0.009)
NAT	0.011	0.011	0.012	0.012	0.012	0.012	0.685	**0.135**	0.011	0.011	0.011	0.013	0.011	0.012	0.013	0.011	0.012
	(0.010)	(0.011)	(0.012)	(0.011)	(0.012)	(0.012)	(0.02)	(0.033)	(0.011)	(0.01)	(0.011)	(0.013)	(0.011)	(0.012)	(0.012)	(0.011)	(0.011)
CAR	0.010	0.01	0.011	0.010	0.01	0.012	0.014	0.819	0.011	0.01	0.009	0.012	0.011	0.012	0.01	0.011	0.012
	(0.010)	(0.01)	(0.010)	(0.010)	(0.009)	(0.012)	(0.015)	(0.032)	(0.011)	(0.009)	(0.009)	(0.012)	(0.01)	(0.01)	(0.011)	(0.011)	(0.011)
FOG	0.011	0.011	0.011	0.010	0.011	0.014	0.023	**0.133**	0.678	0.011	0.011	0.012	0.011	0.011	0.011	0.011	0.011
	(0.011)	(0.01)	(0.010)	(0.010)	(0.011)	(0.012)	(0.027)	(0.035)	(0.011)	(0.010)	(0.01)	(0.011)	(0.011)	(0.011)	(0.01)	(0.011)	(0.01)
EUM	0.019	0.018	0.018	0.020	0.019	0.017	0.016	0.018	0.019	0.686	0.018	0.018	0.019	0.024	0.018	0.029	0.018
	(0.018)	(0.016)	(0.018)	(0.019)	(0.017)	(0.017)	(0.016)	(0.018)	(0.018)	(0.018)	(0.017)	(0.016)	(0.017)	(0.021)	(0.017)	(0.023)	(0.016)
SEI	0.018	0.016	0.016	0.016	0.015	0.016	0.015	0.016	0.017	0.015	0.684	0.015	0.015	0.031	0.016	0.025	0.047
	(0.018)	(0.015)	(0.015)	(0.015)	(0.014)	(0.016)	(0.015)	(0.015)	(0.019)	(0.014)	(0.016)	(0.015)	(0.014)	(0.025)	(0.015)	(0.022)	(0.027)
IJU	0.012	0.012	0.012	0.012	0.012	0.011	0.012	0.018	0.012	0.011	0.011	0.799	0.012	0.011	0.011	0.012	0.015
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.01)	(0.012)	(0.015)	(0.010)	(0.01)	(0.011)	(0.032)	(0.012)	(0.011)	(0.01)	(0.012)	(0.014)
VER	0.014	0.014	0.014	0.014	0.014	0.015	0.014	0.014	0.014	0.015	0.016	0.015	0.753	0.02	0.016	0.016	0.015
	(0.013)	(0.014)	(0.013)	(0.014)	(0.013)	(0.015)	(0.013)	(0.014)	(0.014)	(0.015)	(0.016)	(0.014)	(0.032)	(0.019)	(0.015)	(0.015)	(0.014)
BAK	0.017	0.016	0.015	0.017	0.015	0.015	0.018	0.016	0.016	0.018	0.018	0.015	0.017	0.712	0.016	0.035	0.018
	(0.016)	(0.016)	(0.014)	(0.017)	(0.014)	(0.015)	(0.017)	(0.015)	(0.015)	(0.016)	(0.017)	(0.015)	(0.016)	(0.032)	(0.016)	(0.031)	(0.016)
LIE	0.019	0.018	0.018	0.018	0.017	0.019	0.018	0.018	0.017	0.018	0.018	0.017	0.019	0.024	0.685	0.03	0.019
	(0.017)	(0.017)	(0.017)	(0.017)	(0.016)	(0.018)	(0.018)	(0.017)	(0.015)	(0.017)	(0.017)	(0.015)	(0.017)	(0.025)	(0.018)	(0.023)	(0.018)
NUE	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.017	0.009	0.009	0.009	0.009	0.009	0.01	0.009	0.796	0.054
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.014)	(0.009)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)	(0.029)	(0.024)
CUN	0.009	0.008	0.008	0.008	0.007	0.007	0.007	0.035	0.008	0.008	0.009	0.01	0.008	0.009	0.008	0.014	0.8325(
	(0.009)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.025)	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)	(0.009)	(0.007)	(0.013)	0.035)

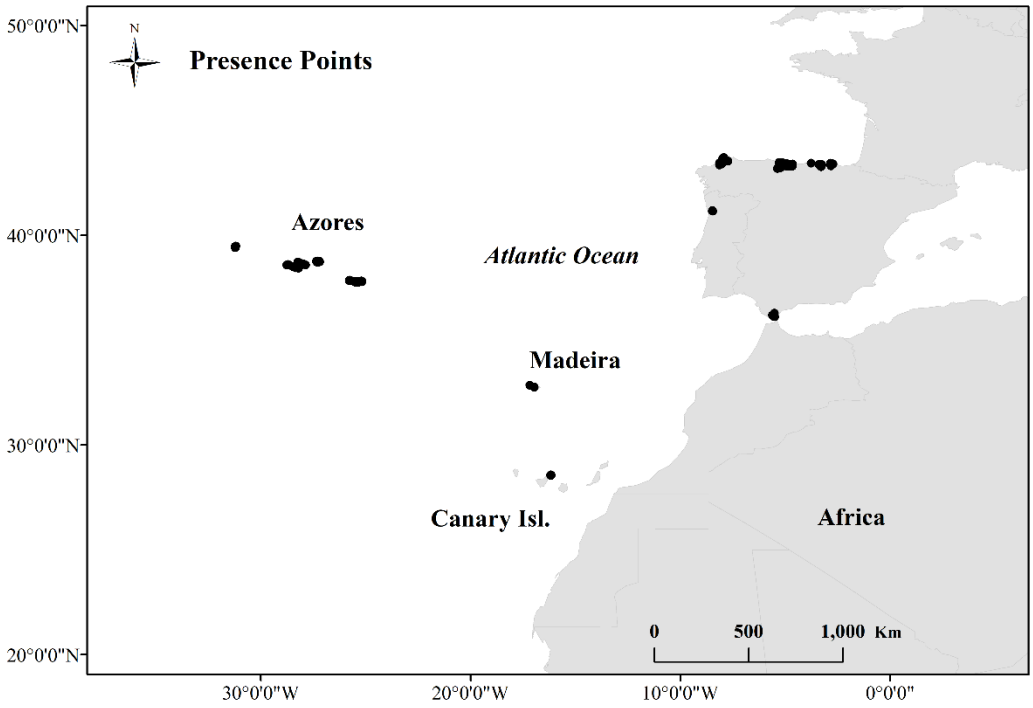


Figure S1. Location of presence records (black dots) used for species distribution modelling (SDM).

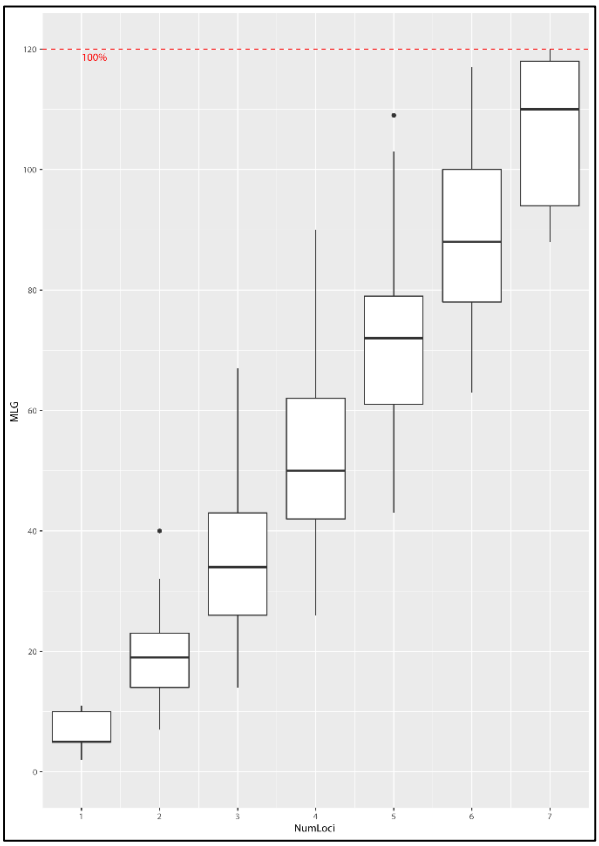


Figure S2. Genotypic accumulation curve showing the resolutive power of the eight microsatellite used in this study.

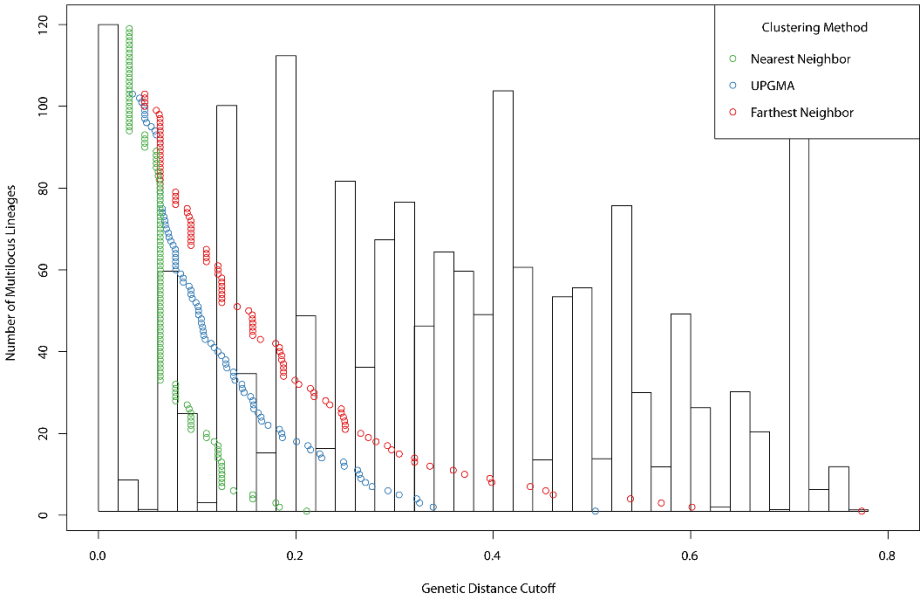


Figure S3. The genetic threshold distance under which two MLGs were considered

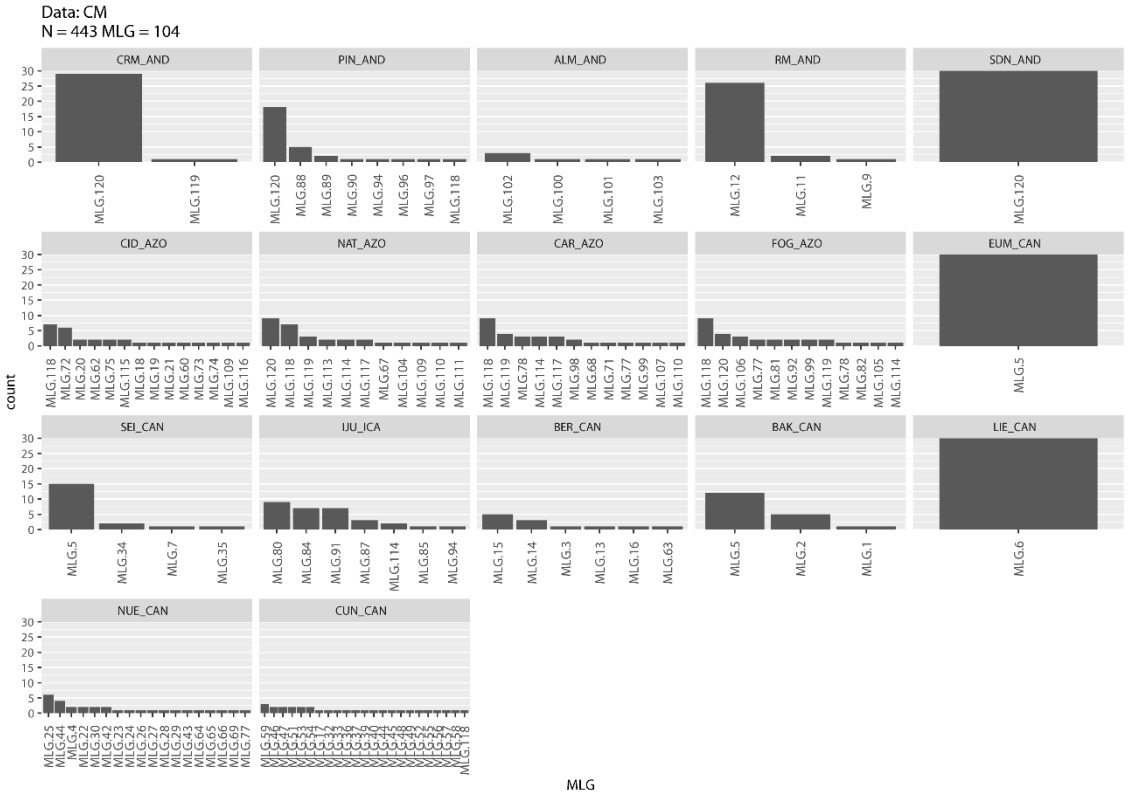


Figure S4a. Distribution of the 104 MLLs among the 130 individuals (genets) across the 17 populations.

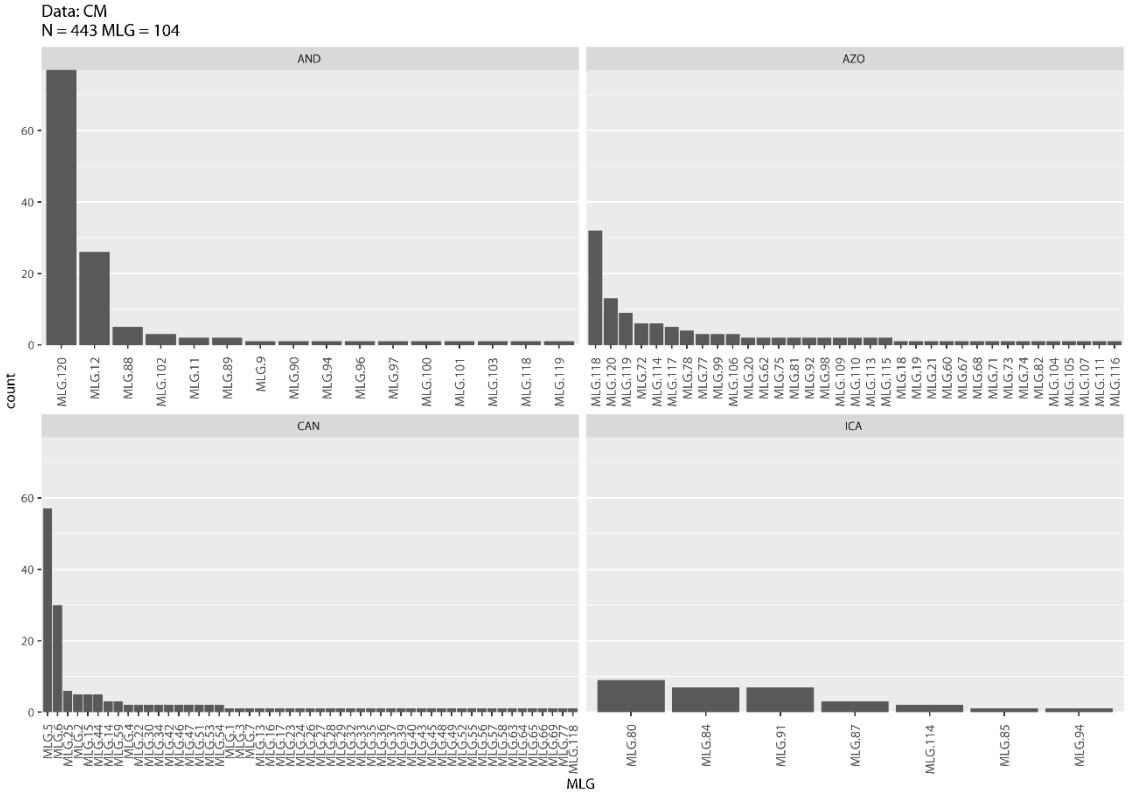


Figure S4b. Distribution of the 104 MLLs among the 130 individuals (genets) across the 4 regions.

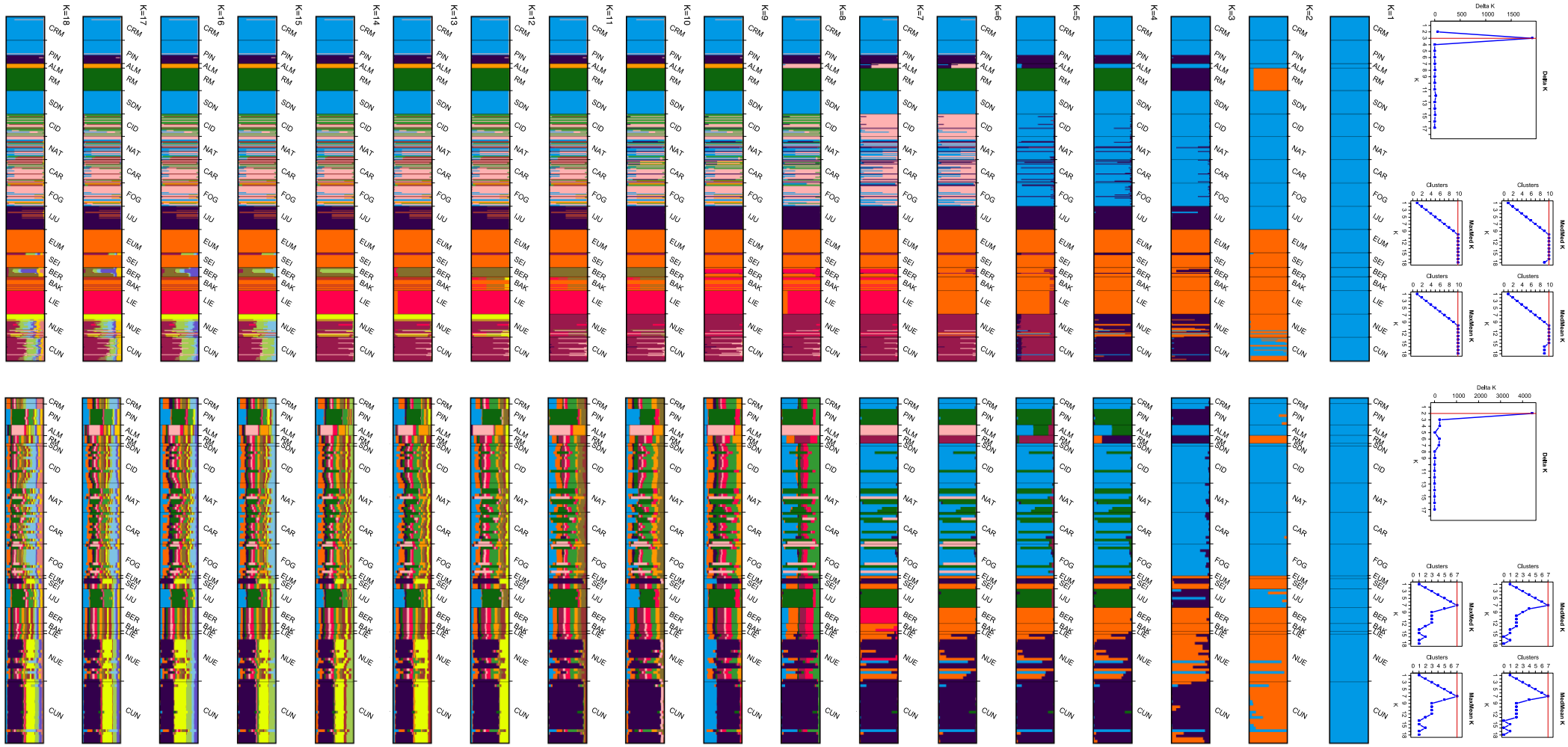


Figure S5. Bar plots showing the STRUCTURE results, using microsatellite data and assuming the non admixture model, when all sampling units were considered (on the left) and when only one individual per MLL per population was used (on the right). Delta *K* method (Evanno et al., 2005) was used to identify the uppermost hierarchical level of genetic structure, while MedMedK, MedMeaK, MaxMedK, and MaxMeaK (with membership coefficient threshold of 0.5; Puechmaille, 2016) were used to identify other levels of genetic partitioning. The *K* values selected by each estimator are shown on the top. For population codes, see Table 1.

Chapter 4. Population variability and differentiation and phylogeographic history in the endemic Ibero-Macaronesian fern *Diplazium caudatum* (Athyriaceae)

Samira Ben-Menni Schuler, Jesús Picazo-Aragonés, Ana Teresa Romero García, Víctor N. Suárez Santiago

Manuscript in preparation

1. Introduction

Macaronesia is considered an exceptional biogeographical and evolutionary model system (Fernández-Palacios et al 2011 and references therein). It comprises the mid-Atlantic, oceanic archipielagos of the Azores, Canaries, Cape Verde, Madeira and Selvagem Islands. The flora of this region presents typical oceanic island characteristics, as a high degree of endemism (20% overall; Humphries, 1979), however is differentiated from these in other aspects, as that they are close to potential continental source areas (García-Talavera, 1997, 1999). This distance between island and mainland was even smaller during glacial periods, perhaps allowing greater and more complex connections between organisms than in other island systems.

It has been thought that island populations should be less variables and more genetically differentiated in relation to continental source populations (Barrett, 1996) due to loss of dispersal capacity (Travis et al., 2010), among other causes. It is assumed the "niche preemption" hypothesis, in relation to island-continent capacity of organisms to colonize new habitats, according to which island colonization could have a higher probability by the available ecological space and conversely, back colonization to the continent be interfered due to interspecific competition (Whittaker & Fernández-Palacios, 2007). Therefore, islands have been considered the "end of the colonization road". However this last could be questioned, especially for organisms that show high dispersal power, as ferns, wich are spore-producing plant; and should be considered the fact that they could return to the continent as a possible biogeographic pattern.

In the case of Macaronesia, and Canary Islands in particular, this back-colonization or even multiple independent colonizations could be a possibility due to the proximity to the continent, the age of the islands (0.8-21 my; Carracedo, 1994) and the existence of potential stepping stones, this is, large volcanic sea mounts between the Canary Islands, Madeira and the continent. Some of this sea mounts are less than 100 m below the sea level and during glacial periods with lower sea level could facilitated a more complex dispersion pattern than only a simple colonization event from continent to island (García-Talavera, 1997). Phylogeographic analyses in Macaronesian non-endemic spore-dispersal plants as mosses and liverworts show the presence of gene flow and several colonizations events both among archipielagos con between those and continents (Vanderpoorten et al., 2008; Freitas & Brehm, 2001).

Engler (1879) proposed the presence of a distinctive endemic element in the Macaronesian flora, different from any other flora in the world, and suggested to be a relict of the widespread tropical flora who lived in Southern Europe during the Tertiary, especially the Palaeogene Period (65-25 Ma), the so-called Palaeotropical geoflora (Mai, 1989, 1991). This flora was comprised of trees of lauroid-shape leaves, palms and ferns dominated this region (Barrón & Peirot, 2006).

Many species belonging to the Tertiary flora saw their area of distribution reduced during the Pliocene and Pleistocene (late Tertiary, 5 million years (ma), and early Quaternary, 2.6 ma) due to the decrease in temperature and the increase in aridity that culminated in the Quaternary glaciations (Barrón, 2003, Pulido et al., 2008, Rodríguez-Sánchez and Arroyo, 2008, Postigo Mijarra et al., 2009, Fernández-Palacios et al., 2011). This caused many Tertiary relict species, including ferns, to have a disjunct distribution today. These relicts are localized in refuges with similar microclimatic conditions to those of that time; these refuges are present along the european atlantic coast and Macaronesia (Jermy, 1984; Cronk, 1992; Vogel et al., 1999b; Quintanilla et al., 2007; Vanderpoorten et al., 2007; Calleja et al., 2009; Liu & Schneider, 2013). Within Europe, the Iberian Peninsula was the last refugium of this Palaeotropical geoflora, although impoverished (Postigo Mijarra et al., 2009), and the major refugia of this flora in this region are the Macaronesian archipielagos (Fernández-Palacios et al., 2011) because of the less impact of the climate deterioration. It is of great interest to reconstruct the evolutionary history of these species and identify unique lineages for their conservation by studying the biogeographic pattern of the different lineages, their population dynamics and the putative past geological-climatic events that could have affected the actual distribution and diversity of them; and phylogeographic aproaches permit generate this knowledge (see Provan & Bennett, 2008).

The relictual character of this flora could be due to the characteristic climate which predominates in Macaronesian archipielagos, and therefore, subtropical taxa in Macaronesia could have persisted and subsequently carry out post-glacial colonizations to the continent.

Diplazium is a genus that belongs to the family Athyriaceae, which groups 350-400 species (Tryon and Tryon, 1982, Rothfels et al., 2012, Wei et al., 2013) with pantropical distribution. Asia is the most diverse region and the possible center of origin of the

genus, which is estimated to have begun to diversify approximately 41.7 ma (Wei et al., 2015). In Europe there are two species of *Diplazium*: *D. sibiricum*, in the Palearctic region in Finland (Wei et al., 2013, 2015), and *D. caudatum*, an European-Macaronesian endemism, Both species are supposed Boreotropical remnants resulting from the decrease in temperature at the end of the Tertiary (Wei et al., 2015).

Diplazium caudatum grows in undergrowth of laurisilva and dark ravines, on very humid soils, with great permanent atmospheric humidity and warm and constant temperatures (Blanca, 1999, Delgado and Plaza, 2006, Moya et al., 2010). Currently it is found, as a Tertiary relict (Barrón & Peyrot 2006), in areas considered as refuges during the glacial periods of the Quaternary (Fernández-Palacios et al 2011), in the Sierras de Algeciras (Cádiz, southern Iberian Peninsula) and Macaronesian Islands (Azores, Cape Verde (only in Santo Atâo), Canary Islands, and Madeira (Blanca, 1999, Delgado and Plaza, 2006; Marrero & Sánchez, 1992). These are regions with subtropical climate and ideal conditions for the development of this species.

No previous studies about the genetic diversity or phylogeography in *D. caudatum* have been made. Dilute the phylogeographic pattern of D. caudatum, a fern that is expected to be a Tertiary relict of the European subtropical flora and present a high dispersion capacity over long distances, can shed light on the complex biogeographic and colonization patterns that could present this flora. This is especially interesting in *D. caudatum*, present in all Macaronesian archipielagos less Selvagem Islands and the South of the Iberian Peninsula, being able to be in the continent due both to the refuge during the glaciations or a recent recolonization.

The purpose of this study is to carry out an exhaustive analysis of the genetic situation and phylogeographic pattern of this Tertiary relict fern with disjunct distribution. The aims, using species-specific microsatellite, one plastid marker, and species distribution modelling (SDM) were: a) to test for the Tertiary relict hypotesis; b) to determine the levels of genetic diversity of *D. caudatum*, as well as detect if there is a structuring of this diversity and analyze how it is structured; c) to establish the evolutionary relationships between the different populations/genetic lineages of *D. caudatum* and to know the geological-climatic events and the population processes that explain their distribution; and d) to know how climate change could affect the distribution of this species in the future.

2. Material and Methods

2.1. Plant Material

Samples of *D. caudatum* were obtained from 12 populations in 4 geographic regions across its distribution range: Andalusia, Azores, the Canary Islands and Madeira. The number of populations per region was between two and four including between 17-30 individuals per population (335 in total; Table 1; Fig. 1).

Table 1. Sampling details of *D. caudatum* populations used in the present study

Code	Location	Voucher	Geographical coordinates	Sample size	
				Microsatellites	ptDNA
Andalusia					
COQ	Canuto de Ojén Quesada		N36.127°/W5.585°	17	5
CRM	Cabecera del río de la Miel		N36.105°/W5.528°	18	5
Azores					
AGU	Agual (Terceira)		N32.734°/W16,886°	30	5
CID	Sete Cidades (São Miguel)		N37.835°/W25.788°	30	5
FUR	Furnas (São Miguel)		N28,134°/W17,273°	30	5
SER	Serreta (Terceira)		N28.119°/W17.224°	30	5
Canary Isl.					
ANC	Ancón Negro (La Gomera)		N28.134°/W17.273°	30	5
CED	Cedro (La Gomera)		N28.120°/W17.225°	30	5
IJU	Ijuana (Tenerife)		N28.560°/W16.172°	30	5
PIJ	El Pijaral (Tenerife)		N28.553°/W16.188°	30	5
Madeira					
FRI	Ribeiro Frio		N32.734°/W16.886°	30	5
POR	Levada Portadela		N32.747°/W16.823°	30	5

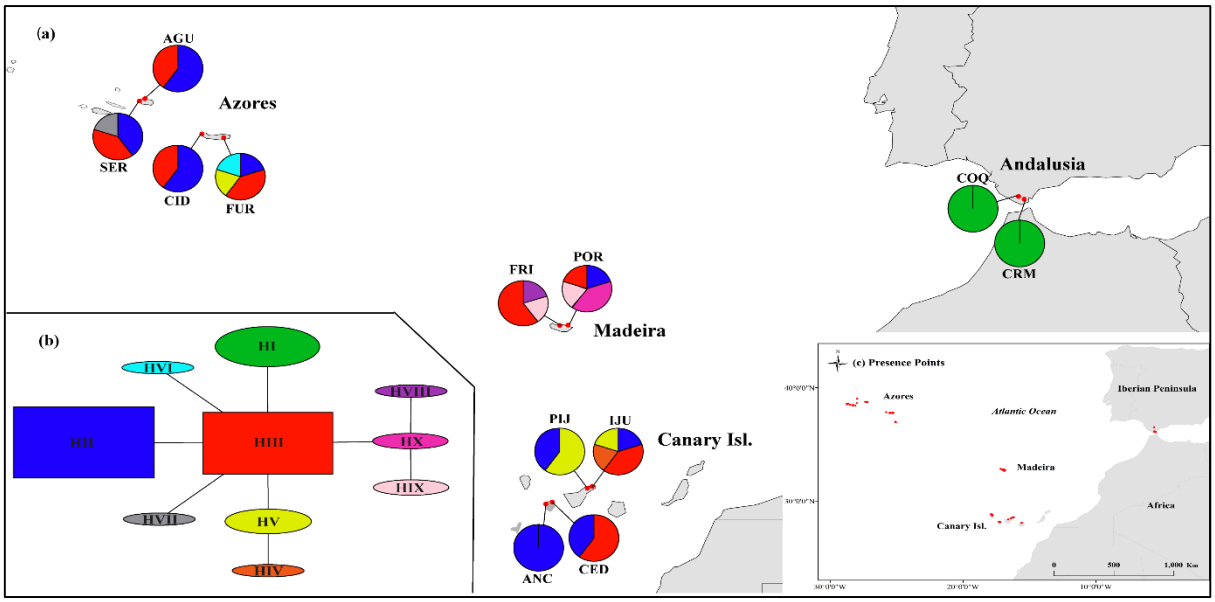


Figure 1. Geographical distribution of haplotypes in the sampled regions and network of haplotypes. a) Geographical distribution of haplotypes and frequency within each sampled population. b) Network representing the phylogenetic relationships of the haplotypes of the plastidial DNA of *D. caudatum* performed by the statistical parsimony method with TCS (Templeton et al, 1992). c) Presence points used for SDM.

2.2. DNA Extraction, Microsatellite Genotyping, and cpDNA Sequencing

For microsatellite analyses, total genomic DNA of the 335 individuals was extracted from about 0.1 g of silica dryed fronds following a modified low-salt CTAB extraction protocol (Arseneau et al 2017). All individuals were genotyped for eight recently developed microsatellite loci described in Ben-Menni Schuler et al. (in prep.), and following the PCR conditions explained in that study. Genotyping was peformed on an ABI PRISM® 3100-Avant Genetic Analyzer (Applied Biosystems, Foster City, CA, United States). Alleles were scored using GENEMARKER v1.85 (SoftGenetics, State College, PA, United States).

For plastid DNA (ptDNA) analysis, after the preliminary search of variable regions (*trn*S-*trn*G, *trn*H-*psb*A, trnLF, *rpl*32-*trn*L*(UAG))*, the spacer region *trn*S-*trn*G was selected as a variable marker, and a partial sequence (toward 5’ end) of the intergenic spacer was amplified and sequenced for a subsample of 5 individuals per population (60 individuals in total). We could not use the entire region due to the presence of long mononucleotide runs, from about 16 base pairs (bp), which prevented sequencing. The specific primers DC-trnS-GCU2: 5’-ATT AGC AAT CCG ACG CTT TA–3’ and DC-

trnG-UCC2: 5’-ATT CGA ACC CGC ATC AGT AG- 3’, designed from the general ones (Shaw et al., 2005), were used to amplify the entire *trn*S-*tr*nG region in all individuals. Then, the internal primer 5'-trnG2S (Shaw et al., 2005) was used to sequencing the 5’ fragment.

PCR reactions were performed in 25-µL reactions containing 50 ng of genomic DNA, 1.25 µM of each primer, and 12.5 µL of the Kapa 2G Robust HotStart ReadyMix (Kapa Biosystems, Massachusetts, USA). Cycling parameters were those described in Shaw et al. (2005). Sequencing was performed on an ABI PRISM® 3100-Avant Genetic Analyzer (Applied Biosystems, Foster City, CA, United States). The resulting sequences were aligned using the Clustal algorithm in the sequence editor BIOEDIT v7.0.5.3 (Hall, 1999), and then adjusted by eye.

2.3. Genetic Diversity and Structure

2.3.1. Microsatellites

In order to infer the clonal identity of the sampling units (all individuals sampled), this is, the prevalence of asexual reproducton in *D. caudatum*, we tested first, the resolutive power of the eight microsatellite markers on R package Poppr (Kamvar et al., 2014). Secondly, we calculated the probability of finding the same multilocus genotype (MLG) due to a sexual event (*p*sex), this is, that have originated from different reproductive events and belong to different genets, based on the observed allele frequencies, the sample size of the data set, taking into account departures from Hardy-Weinberg equilibrium in the population following Arnaud-Haond et al. (2007) and using Fis to calculate genotipic frequencies (*p*gen); calculated using MLGsim v 2.0 (Stenberg, Lundmark & Saura, 2003). We ran 1.000 simulations to obtain the distribution of simulated *p*sex values with which to contrast the significance of real *p*sex. Finally, to define the existence of multiple locus lineages (MLL), this is, to check if different MLGs belong to the same clone or clonal lineage we analized the distribution of the frequencies of genetic distances between pairs of MLGs, with the function mlg.filter and using Bruvo distances on R package Poppr (Kamvar et al., 2014). The genetic threshold distance under two MLGs belonging to the same MLL were considered was estimated using the farthest neighbor method.

The descriptors of clonal diversity were calculated with the function poppr on R package POPPR (Kamvar et al., 2014) as follows: a) to characterize the clonal richness we calculated: number of MLLs, number of expected MLLs (eMLLs) and clonal richness (R) corrected for the sampling size. b) To characterize the genotypic diversity we calculated: the Simpson's index (lambda; corrected for the sampling size) and the clonal eveness index (E.5), which is an indicator of how equally each MLL is represented. Finally, to test the predominant reproductive model we calculated the standarized association index ($\overline{r}$ D). This descriptors were estimated for all sampling units (except $\overline{r}$ D calculated also for MLLs) and both per population and region.

The descriptors of genetic diversity were calculated as follows: a) the allelic richness (Ar) on R package HIERFSTAT version 0.4-22 (Goudet 2005) with the function allelic.richness and b) the observed and expected heterocigosity (HO and HE, respectively) and fixation index FIS using GenoDive (Meirmans & Van Tienderen, 2004). Furthermore, in order to test the effect of asexual reproduction on intrapopulation genetic diversity (Ar, HO, HE and FIS), we calculated genetic descriptors including both all sampling units and only single representing genotype from each clonal lineage (MLL) per population, and Student's *t* tests were performed to explore significant differences between data sets (including vs. excluding clones). All descriptors were calculated both per population and region.

2.3.2. ptDNA

Genetic diversity was assessed by the number of haplotypes (*ha*), haplotype diversity (*Hd*), and nucleotide diversity (π) calculated using ARLEQUIN v3.5.2.2 (Excoffier & Lischer, 2010). All diversity indices were calculated both per population and region.

2.4. Genetic structure and phylogeography

Standard and hierarchical analyses of molecular variance (AMOVA; Excoffier, Smouse & Quattro, 1992) were used to test for intraindividual diversity, this is, random mating and to evaluate the distribution of genetic variability; this last within and among populations, and among the 4 geographical regions. For microsatellites, this analysis was made with all sampling units (including clones) and with only MLLs on R package

POPPR using the function poppr.amova, and for signification test the function randtest, running 1000 replicates; with cpDNA markers we used the program ARLEQUIN and the significance was tested with permutation tests (10.000 permutations). Population genetic structure was analysed using different approaches with our microsatellite data. First, pairwise F_{ST} values were calculated with both all sampling units and with only MLLs among populations using GenoDive; the significance of F_{ST} was tested by a permutation test with 10 000 permutations. Second, the Bayesian algorithm implemented in STRUCTURE v2.3.4 (Pritchard et al., 2000) was used to evaluate the number of genetic clusters (*K*) with all sampling units The number of clusters tested ranged from one to 18, with 10 replicates per *K*, using the no admixture model and independent allele frequencies. The burn-in period and Markov Chain Monte Carlo (MCMC) iterations were set to 50,000 and 10^6, respectively. The optimal number of clusters was estimated with the online tool STRUCTURESELECTOR (Li and Liu, 2018). We identified the uppermost hierarchical level of genetic structure using the delta *K*-method (ΔK; Evanno et al., 2005), which accurately identifies it when the populations are evenly sampled (Puechmaille, 2016). To explore other levels of genetic partitioning, we used the four independent estimators proposed by Puechmaille (2016; MedMedK, MedMeaK, MaxMedK, and MaxMeaK) considering a membership coefficient threshold of 0.5. To align and visualize the STRUCTURE output across the 10 replicates, we used the online program CLUMPAK (Kopelman et al., 2015). Third, the genetic structure was also assessed using a model-free multivariate statistics-based clustering method, a discriminant analysis of principal components (DAPC) on R package ADEGENET (Jombart et al., 2010) with all sampling units The function *xvalDapc* from ADEGENET was used to select by crossvalidation the correct number of principal components with 1,000 replicates using a training set of 90% of the data. The number of principal components was chosen based on the criteria that it had to produce the highest average percentage of successful reassignment and lowest root mean squared error (Jombart et al., 2010). A Minimun Spanning Network (MSN) was made with all sampling units on R package POPPR using the Bruvo distances and the function bruvo.msn.

For cpDNA, an haplotype network was reconstructed following the statistical parsimony method (Templeton et al., 1992) as implemented in TCS v1.21 (Clement et al., 2000).

2.5. Gene Flow and Demographic Analyses

We tested the connectivity among populations by estimating the migration rates among them, in all cases with all sampling units (including clones). Thus, to know whether there was recent (over two to three generations) gene flow between the populations, we estimated migration rates (*m*) between all individual populations using a Bayesian assignment test with the software BAYESASS v1.3 (Wilson and Rannala, 2003). As program settings, the default values were used (MCMC iterations, 3×10^6; length of the burn-in, 999,999; sampling frequency, 2,000; delta value, 0.15). Isolation by distance (IBD) was tested fir the 17 populations using regression of pairwise F'_{ST} distances [determined with GenoDive using them transformed as $F_{ST}/(1 - F_{ST})$] and logarithms of geographical distances between populations, using a Mantel test in GenoDive. Neutrality tests were carried out with cpDNA, Fu's *F* (Fu, 1997) and Tajima's *D* (Tajima, 1989), to detect possible historical demographic processes (expansion or contraction), using ARLEQUIN and cpDNA. Both tests were performed considering populations and geographical regions. The level of significance of both statistics was obtained by 1,000 simulated samples. In addition, cpDNA sequences were used to test for evidence of population size fluctuations between *D. caudatum* regions by constructing Bayesian Skyline Plots with BEAST (BSP; Drummond *et al*., 2005; see Methods S1 for details).

2.6. Haplotype phylogeny and Dating

Phylogenetic relationships among cpDNA haplotypes of *D. caudatum* and the outgroup species were inferred using Bayesian Inference (BI), with MrBayes v3.1.2 (Ronquist et al., 2012**).** The analysis was run using the selected substitution model, Hasegawa-Kishino-Yano, with proportion of invariable sites (HKY + I; Hasegawa et al., 1985) determined by jModeltest v2.1.10 (Posada, 2008) and included 2 million generations with 2 simultaneous runs (8 Markov Monte chains Carlo, MCMC), starting from random trees that were sampled every 100 generations. The TRACER v1.7 program (Rambaut et al., 2018) was used to visualize the results and confirm the convergence and stationarity of the races. 25% of the initial trees, from the pre-convergent phase (burn-in phase) were removed. The rest of the trees were used to construct the phylogenetic consensus tree ("50% majority rule consensus").

To date divergence events between plastidial haplotypes we used BEAST v2.0 computer program (Drummond et al., 2014). The divergence times were obtained using the selected HKY + I substitution model, a strict molecular clock obtained by PAML v4 (Yang, 2007) and a background tree ("priors") according to the "Calibrated Yule Model". The nodes of the tree were constricted according to the phylogeny obtained with MrBayes and calibrated using a normal distribution, establishing the confidence interval, for the estimated divergence time for *D. caudatum* according to Wei et al. (2015). The length of the Markov chains was established at 10 million generations, sampling the trees every 1,000 generations. All these parameters were established with the BEAUti program included in BEAST. The TRACER v1.7 program (Rambaut et al., 2018) was used to visualize the results obtained and confirm the convergence of the chains and eliminate the sampled trees during the pre-convergent phase ("burn-in"). The initial 10% of trees sampled in each race were eliminated. The trees obtained were joined in a tree of maximum credibility of the clade ("maximum clade credibility tree") using TREEANNOTATOR, also included in BEAST, and was visualized using the tree editor FIGTREE v1.4.3 (Rambaut, 2016).

2.7. Species Distribution Modelling

To identify potential refugial and future distribution areas for *D. caudatum*, species distribution modelling (SDM) was performed. This analysis requires presence occurrence data of the studying species and environmental variables. As environmental data we used 19 BIOCLIM variables at a resolution of 2.5 arc-minutes (ca. 5 km) representing different time periods during past, present and future climatic conditions. Past and current climate data were available from the WorldClim database (www.worldclim.org; Hijmans *et al*., 2005) and included data for the current-day period (1,950– 2,000), the Last Glacial Maximum (LGM; *c*. 21 ka) simulated by CCSM model (the Community Climate System Model), and for the Last Interglacial period (LIG; *c*. 120ka). We obtained predictions for future climatic conditions in 2080 for the most impacting IPCC's climate scenario: RCP8.5 (Van Vuuren et al., 2011) available through the CCAFS Climate portal (www.ccafs-climate.org). Soil data were obtained from SoilGrids.org (Hengl et al. 2014) but was not used with past climatic conditions because of the lack of this type of maps. Highly correlated variables (Pearson's $R \geq 0.8$) were reduced to eight uncorrelated variables (Table S1) used as predictors to calibrate the distribution models. Species occurrence data were obtained from a collection of references in databases (the Global

Biodiversity Information Facility data portal (http://www.gbif.org/), the Biodiversity databank of the Canary Islands (http://www.biodiversidadcanarias.es/atlantis/common/index.jsf), and the Azores Biodiversity databank (http://www.atlantis.angra.uac.pt/atlantis/common/index.jsf), literature (Benl, 1971; Queirós et al 1991), and our own field records. A total of 323 presence records were finally compiled (Fig.1). To perform the SDM we applied Maximum Entropy Modelling implemented in the software package MAXENT 3.4.1 (Phillips *et al*., 2006). Models were generated using cross-validation of 5 replicate runs. Model performance was assessed based on the area under the receiver operating characteristic curve (AUC). The contribution of each predictor variable in the model was analysed by the permutation importance and percent contribution coefficients (Table S1). A final reduced model including the most important variables (Mean Diurnal Range and Minimum Temperature of Coldest Month), was finally computed (Martínez, Viejo, Carreño, & Aranda, 2012).

3. Results

3.1. Genetic diversity and clonality

3.1.1. Microsatellites

In relation to the resolutive power of the eight microsatellite used in this study, with seven loci almost 100% of the MLG are resolved as shown in the genotypic accumulation curve (Fig. S1). A total of 294 multilocus genotypes (MLG) were found among the 335 individuals sampled and 282 multilocus lineages (MLL); however, eight MLLs are distributed among individuals of different populations, therefore, we kept these with their original MLGs resulting finally in 290 MLLs (Table 2). In general, *D. caudatum* shows low prevalence of clonality (clones defined as individuals sharing the same MLL), showing similar values of clonal richness and genotype diversity in populations and regions of Macaronesia: Azores, the Canary Islands and Madeira, being the last one the most diverse (34.66 eMLLs, R = 0.983, lambda = 0.999 and $\overline{r}$ D = 0.007 although not significative). However, Andalusia contrasts with the other regions (11 eMLLs, R = 0.294, lambda = 0.798 and $\overline{r}$ D = 0.334), indicating a high presence of clones between the individuals of populations COQ and especially CRM (Table 2).

Table 2. Intrapopulation diversity in *D. caudatum*. Clonality descriptors were separated into clonal richness and genotype diversity. *N*, number of individuals; MLL, number of different multilocus lineages; eMLL, expected number of different multilocus lineages; R, corrected clonal ratio; lambda, Shannon corrected index; E.5, clonal eveness value; $\overline{r}$ D, association index including all sampling individuals; $\overline{r}$ Dc, association index including only unique MLLs (* $P < 0.05$). N.A.: not applicable

		Clonal richness			Genotype diversity			
Population	*N*	MLL	eMLL	R	lambda	E.5	$\overline{r}$ D	$\overline{r}$ Dc
Andalusia	35	11	11	0.294	0.798	0.587	0.334*	0.103*
COQ	17	9	9	0.500	0.912	0.875	0.107*	0.047
CRM	18	2	2	0.059	0.294	0.676	NA	NA
Azores	120	107	33.3	0.883	0.996	0.845	0.015	-0.002
AGU	30	30	17	1	1	1	-0.018	-0.018
CID	30	28	16.23	0.931	0.993	0.927	-0.019	-0.028
FUR	30	26	15.6	0.862	0.989	0.915	0.080*	0.054*
SER	30	26	15.6	0.862	0.989	0.915	0.047	0.016
Canary Isl.	120	114	34.44	0.949	0.999	0.964	0.021*	0.019*
ANC	30	29	16.7	0.965	0.998	0.981	0.019	0.014
CED	30	27	15.9	0.896	0.991	0.92	0.112*	0.088*
IJU	30	29	16.7	0.965	0.998	0.981	0.007	0.002
PIJ	30	29	16.7	0.965	0.998	0.981	0.017	0.005
Madeira	60	59	34.66	0.983	0.999	0.990	0.007	0.003
FRI	30	30	17	1	1	1	0.018	0.018
POR	30	29	16.7	0.965	0.998	0.981	0.022	0.008
Total	335	290	16.67	0.865	0.997	0.682	0.104	0.099

101 alleles were observed from the eight loci surveyed (average per locus: 12.625). The total genetic diversity (H_E) for *D. caudatum* was 0.495 (Table 3). At the population level, allelic richness ranged from 1.125 in CRM to 7.282 in FRI. The lowest value of genetic diversity was found in CRM (H_E = 0.083), whereas the highest value was found in FRI (H_E = 0.626); the mean genetic diversity value per population was 0.511 (Table 2). CRM population showed a high proportion of monomorphic loci (seven of the eight analyzed). At the geographical region level, the most diverse region was Madeira and the less diverse Andalusia (Table 3). Both Azores and the Canary Islands populations and regions show similar genetic diversity values. Population exclusive alleles (31 in

total) were detected in populations of all regions less Andalusia, and 23 of the 31 were present in populations from Madeira (FRI, 12; POR, 11; Madeira in general, 14). All populations showed statistically significant positive values of inbreeding coefficients, except CRM and CID populations in which statistical significance was not reached.; these values are low and similar in all Macaronesian archipielagos, but high in Andalusia. Comparing diversity descriptors when including all sampling individuals or only MLLs, *Ar* decreases, especially in Madeira and H_O, H_E and F_{IS} increase slightly or remains similar without clones.

3.1.2. ptDNA

Plastid DNA sequence alignment included 60 sequences in total, with 722 base pairs (bp) in length and it included nine variable positions. The total number of haplotypes found was 10. Results for the diversity indices are shown in Table 3. At the population level, FRI is the most diverse population and both the Andalusian populations and ANC (Azores) show null diversity values. At the geographical region level, the most diverse region was Madeira and the less diverse, Andalusia (Table 3).

Table 3. Genetic descriptors for microsatellite and plastid data in the populations of *D. caudatum* studied. For microsatellite, descriptors (Ar, $H_{O,}$ H_E and F_{IS}) are calculated when including all individuals sampled per population and those including unique genotypes without clones. *N*, number of individuals; Priv, number of private alleles; *Ar*, allelic richness; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{IS}, inbreeding coefficient (* $P < 0.001$). *ha*: number of haplotypes; *Hd*: haplotype diversity; π: nucleotide diversity

		Microsatellites									**cpDNA**			
Population	*N*	Priv	*Ar*	Ar 1 MLL	H_O	H_O 1 MLL	H_E	H_E 1 MLL	F_{IS}	F_{IS} 1 MLL	*ha*	Priv	*Hd*	π
Andalusia	35	0	1.625	1.625	0.025	0.045	0.228	0.297	0.890*	0.847*	1	1	0.000	0.000
COQ	17	0	1.625	1.478	0.029	0.042	0.251	0.283	0.883*	0.853*	1	0	0.000	0.000
CRM	18	0	1.125	1.125	0.021	0.063	0.020	0.063	-0.063	0.000	1	0	0.000	0.000
Azores	120	4	4.203	3.179	0.237	0.251	0.353	0.369	0.328*	0.319*	5	2	0.663	0.001
AGU	30	0	2.757	1.748	0.263	0.263	0.357	0.357	0.264*	0.264*	2	0	0.600	0.001
CID	30	1	2.871	1.722	0.321	0.335	0.334	0.348	0.041	0.037	2	0	0.600	0.001
FUR	30	2	2.874	1.723	0.171	0.178	0.328	0.347	0.479*	0.488*	4	1	0.900	0.002
SER	30	1	3.202	1.722	0.196	0.221	0.314	0.342	0.377*	0.353*	3	1	0.800	0.001
Canary Isl.	120	4	4.930	3.685	0.224	0.230	0.428	0.428	0.476*	0.463*	4	1	0.679	0.001
ANC	30	0	2.586	1.661	0.200	0.198	0.316	0.318	0.367*	0.377*	1	0	0.000	0.000
CED	30	1	2.679	1.712	0.183	0.204	0.340	0.335	0.460*	0.392*	2	0	0.600	0.001
IJU	30	1	2.840	1.742	0.233	0.233	0.340	0.343	0.314*	0.322*	2	1	0.600	0.002
PIJ	30	2	3.441	1.826	0.279	0.284	0.381	0.386	0.267*	0.263*	4	0	0.900	0.002
Madeira	60	37	9.576	6.514	0.417	0.417	0.630	0.630	0.338*	0.338*	5	3	0.822	0.002
FRI	30	11	7.282	2.567	0.438	0.438	0.626	0.626	0.301*	0.301*	3	1	0.700	0.002
POR	30	12	6.915	2.535	0.396	0.397	0.615	0.617	0.356*	0.357*	4	1	0.900	0.002
Total	335	45			0.227	0.238	0.495	0.505	0.354*	0.362*	10	7		

3.2. Genetic structure and phylogeography

Standard AMOVA showed that the higuest genetic variation is found within samples (48.19%, $P < 0.001$), being the 24.92% of the variation between samples within populations and the 26.89% between populations (both $P<0.001$; Table 4). Considering geographical regions, the variation among them was just of 16.13%, although significant. All types of structuring, standard vs hierarchied and including all sampled individuals vs. including only MLLs resulted in similar values of variation in each hierarchical level and show the higuest genetic variation within samples. Plastid AMOVA results show that the highest percentage of genetic variation resides within the populations (64.41% of the variation), and that the 29.44% is at the interregional level;

only 6.15% of the variation lies in the interpopulation component. The differences between the different regions and within the populations were found to be significant, contrary to what happens at the interpopulation level (Table 4).

Table 4. Hierarchical analysis of molecular variance (AMOVA)

Souce of variation	d.f.	Sum of squares	Percentage of variation	F-value	*P*-value
Microsatellites					
All sampling units					
Without grouping					
Within samples	335	650	48.19	0.518	<0.001
Between samples within populations	323	1274.937	24.92	0.341	<0.001
Between populations	11	706.699	26.89	0.269	<0.001
Total	669	2631.635	100		
Among regions (4 groups)					
Within samples	335	650	46.37	0.536	<0.001
Between samples within populations	323	1274.937	23.98	0.341	<0.001
Between population within regions	8	288.806	13.52	0.161	<0.001
Between regions	3	417.793	16.13	0.161	<0.001
Total	669	2631.636	100		
MLLs					
Without grouping					
Within samples	294	632	53.75	0.462	<0.001
Between samples within populations	282	1194.834	26.09	0.327	<0.001
Between populations	11	476.138	20.16	0.201	<0.001
Total	587	2302.973	100		
Among regions (4 groups)					
Within samples	3	128.039	52.28	0.477	<0.001
Between samples within populations	8	120.433	25.38	0.327	<0.001
Between population within regions	584	920.342	11.93	0.133	<0.001
Between regions	595	1168.814	10.41	0.104	<0.001
Total			100		
cpDNA					
Among regions (4 groups)					
Within populations	48	13.200	64.41		<0.001
Between populations within regions	8	3.250	6.15		0.16618
Between regions	3	6.667	29.44		<0.01

Total	52	23.117	100

At the population level and all sampling units, all pairwise comparisons were significant less FUR-SER (Azores) (Table S2). Azores and Madeira populations show low intra-regional population differentiation. In the Canary Islands the intra-regional values are similar to the inter-regional ones in the comparisons with Azores and Madeira, showing ANC even higher values. In Andalusia, all F_{ST} values were high, and at the interregional level, both populations from this region were the most differentiated, especially CRM, showing this less differentiation with the populations of Madeira. The rest of the pairwise comparisons between populations from different regions showed similar levels of differentiation. Analyses with only MLLs show similar results although all CRM values were not significant (Table S2).

The Bayesian clustering method, as implemented in STRUCTURE, recognized four genetic clusters as the uppermost hierarchical level of genetic partitioning according to the higuest ΔK peak (Fig. 2). These clusters could be identified as the populations from Andalusia (CL1), Madeira (CL2) and Azores and the Canary Islands form CL3 and CL4, although this clusters are not structured by regions, showing many populations influence of both clusters. The results interpreted using the method of Puechmaille (2016) revealed seven clusters ($K = 7$; Fig. 2) as the most likely group structure, considering locations with $Q \geq 0.5$ in any inferred cluster. In this case, CRM and COQ form one cluster, individuals from Azores and the Canary Islands less ANC have similar probabilities of belonging to three clusters less ANC that form a different cluster. Both populations from Madeira form a mixing from the sixth and seventh cluster. COQ shows high presence of the cluster belonging to populations PIJ and IJU from the Canary Islands.

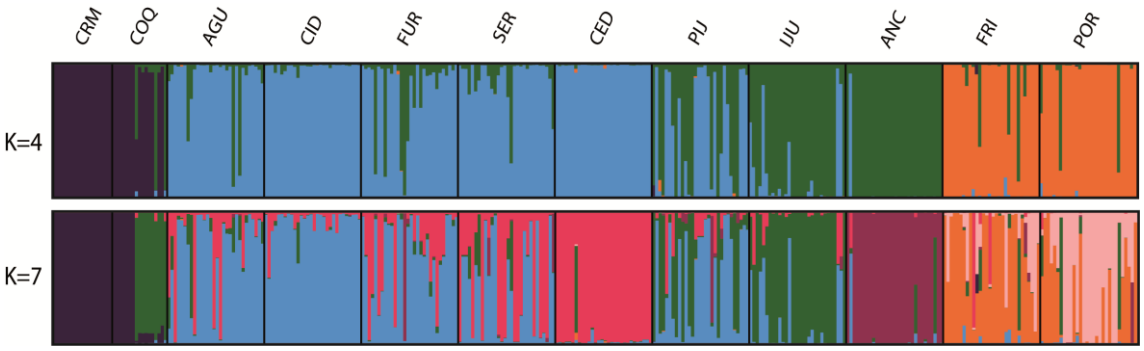


Figure 2. Estimated genetic structure based on microsatellite data using the Bayesian approach implemented in STRUCTURE. Histograms of individual assignment to clusters show the two most probable structuring, *K* = 4 and 7.

Discriminant analysis of principal component yielded in a similar result as the STRUCTURE for *K*=4, although DAPC does not relate COQ with Canarian populations (Fig. 3).

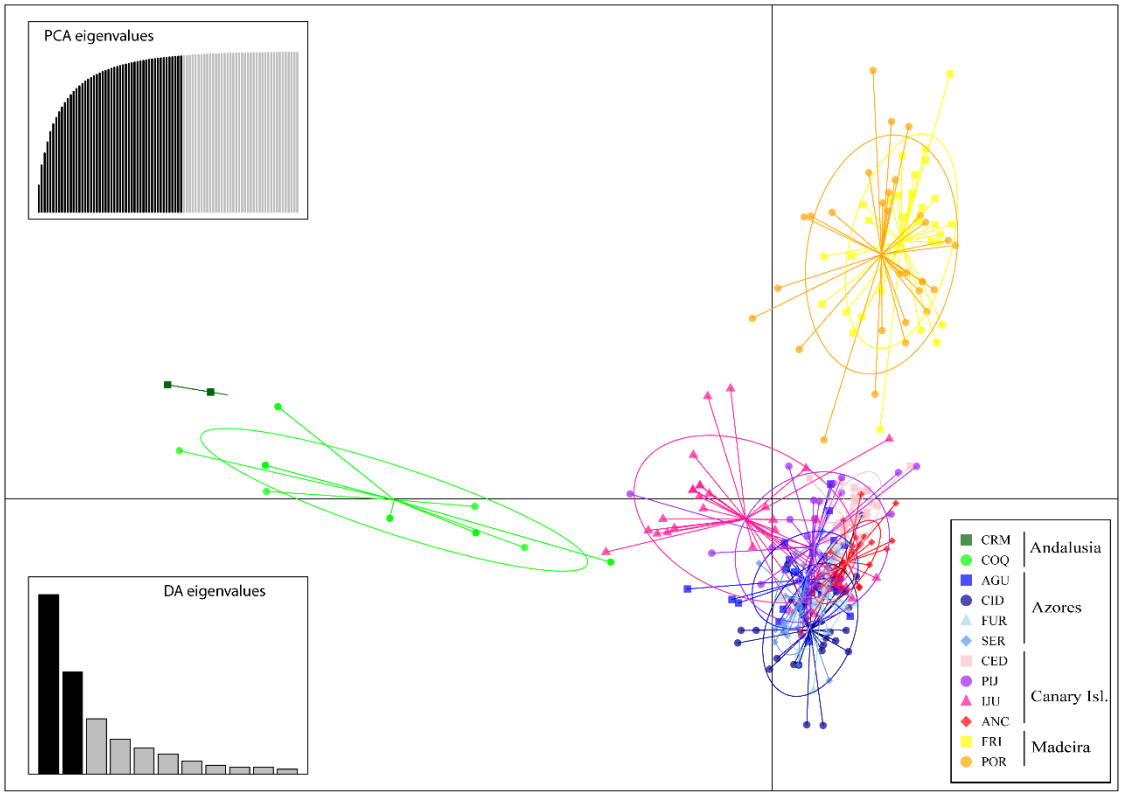


Figure 3. Result of the discriminant analysis of principal components (DAPC) using microsatellites.

The representation on a map of the ptDNA haplotype frequencies and distributions suggests a geographical structuring of them (Figure 1a). Haplotypes H-II and H-III are the most frequent and widespread, being present in all the sampled macaronesian archipelagos. H-I is exclusive from Andalusia, H-V is present in Azores and the Canary

Islands; H-IV is exclusive from the Canary Islands, H-VI and H-VII from Azores and H-VIII, H-IX and H-X from Madeira (Fig.1, Table 3).

3.3. Gene Flow and Demographic Analyses

The results of BAYESASS indicated that current exchange of genes could be occurring from AGU to CID, FUR, SER, PIJ and IJU (migration rate [*m*] = 0.243; 0.224; 0.236; 0.242 and 0.212, respectively; Table S3). Estimates of *m* that did not exceed 0.110 (the upper value of the confidence interval when there is no information in the data) are considered that no current gene flow exist between populations. The Mantel test showed the existence of isolation by distance across the populations ($r = 0.217$, $P < 0.01$) by including all populations, however, when we exclude populations from Andalusia the test was not significant ($r = 0.093$, $P = 0.179$).

Fu's *F* and Tajima's *D* tests resulted not significative except for the *F* in FUR, IJU, and POR populations, being negative, indicating expansion of these populations (Table S4). At the regional level, although not significant, the trend of the results of these tests seems to indicate population expansion from Azores and Madeira (highly negative values). The BSP analyses found evidence for range expansion when we include in the analyses all the distribution range of the species, this starting from around 100 thousand years ago (Ka; beginning of the Würm glaciation), but not for a specific region (Figure 4).

3.4. ptDNA dating and phylogenetic analysis

The phylogenetic tree resulting from the Bayesian analysis with MrBayes (Figure 5) shows how the haplotypes of *D. caudatum* form a monophyletic group, and haplotypes from H-I to H-VII form a clade. Within this clade, relationships between haplotypes are not resolved.

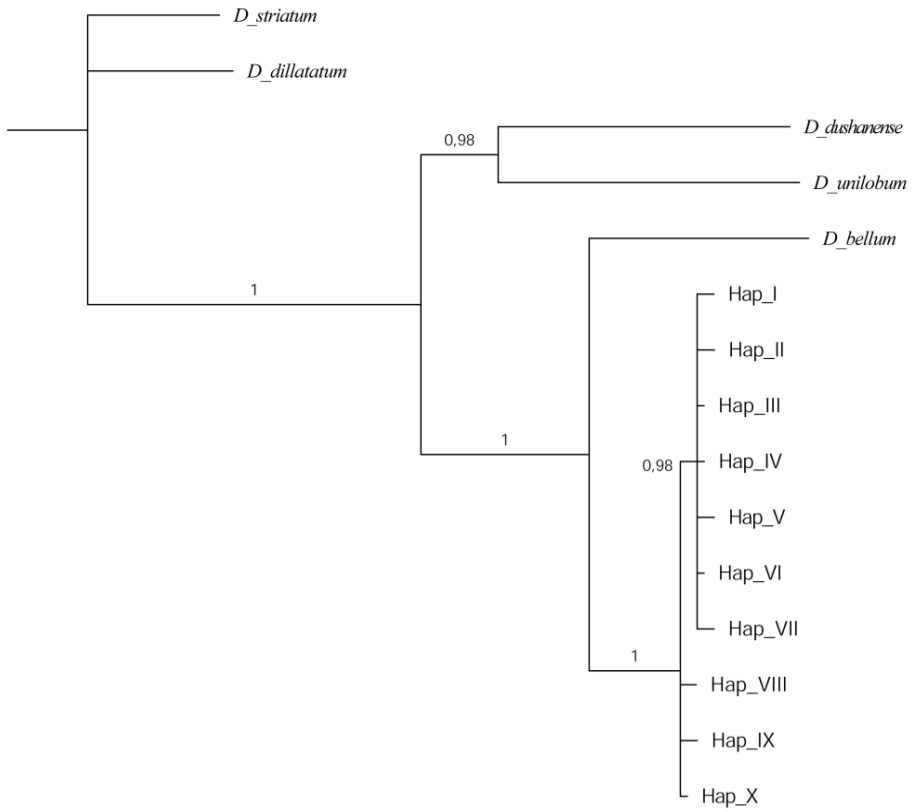


Figure 5. 50% consensus phylogenetic tree "majority rule" obtained from the Bayesian Inference analysis of the trnS-trnG plastidial DNA region of D.caudatum and the external group used. The numbers on the branches indicate the posterior probability of each clade.

The lineage of *D. caudatum*, according to the dating made with BEAST, originated 19.11 [13.54-24.96] mllion years ago (m.a.). The estimated time for the event of initial divergence of the haplotypes of the *trn*S-*trn*G region was 1.82 [0.51-3.35] m.a. coinciding with the end of the first interglacial period and beginning of the second Pleistocene glaciation (Figure 6).

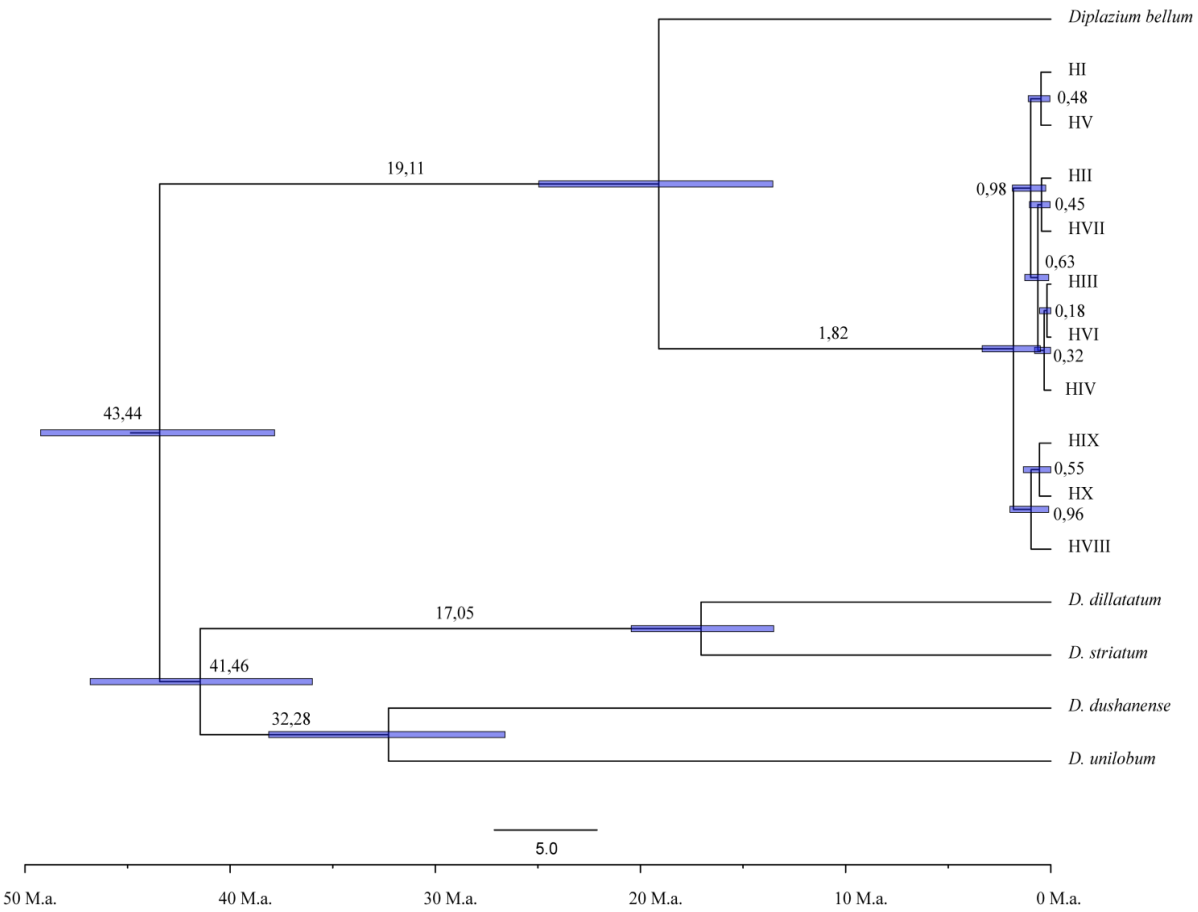


Figure 6. Dendrogram of the 10 haplotypes of the plastidial DNA of *D. caudatum* and the 5 haplotypes of other species of the genus used as an external group in the BEAST analysis. The divergence time (millions of years) of the nodes is shown and 95% of the highest posterior density (HPD) is indicated.

3.5. Species distribution modelling

For all models AUC values were high (minimum value of AUC = 0.992). The predictions obtained by MAXENT for the present and the LIG show similar suitables areas for the habitat of *D. caudatum* in relation to the current distribution of the species, adding the Portuguese coast as a potential area. It is noteworthy that during the LIG, there is less suitability in the south of the Iberian Peninsula in relation to the present (Fig.7). According to the LGM outputs, the shelters of this species were located in the Macaronesian islands, the lower half of the Atlantic coast of the Iberian Peninsula and the coast of Morocco; during this time, more available areas were shown in relation to LIG, less Madeira with the same suitability during both LGM and LIG, and Azores, showing a higuer availability during LIG.. The future projections show a clear decrease

of the suitability in the Iberian Peninsula, especially in Andalusia, staying in the Macaronesian islands with a high probability (Fig.7).

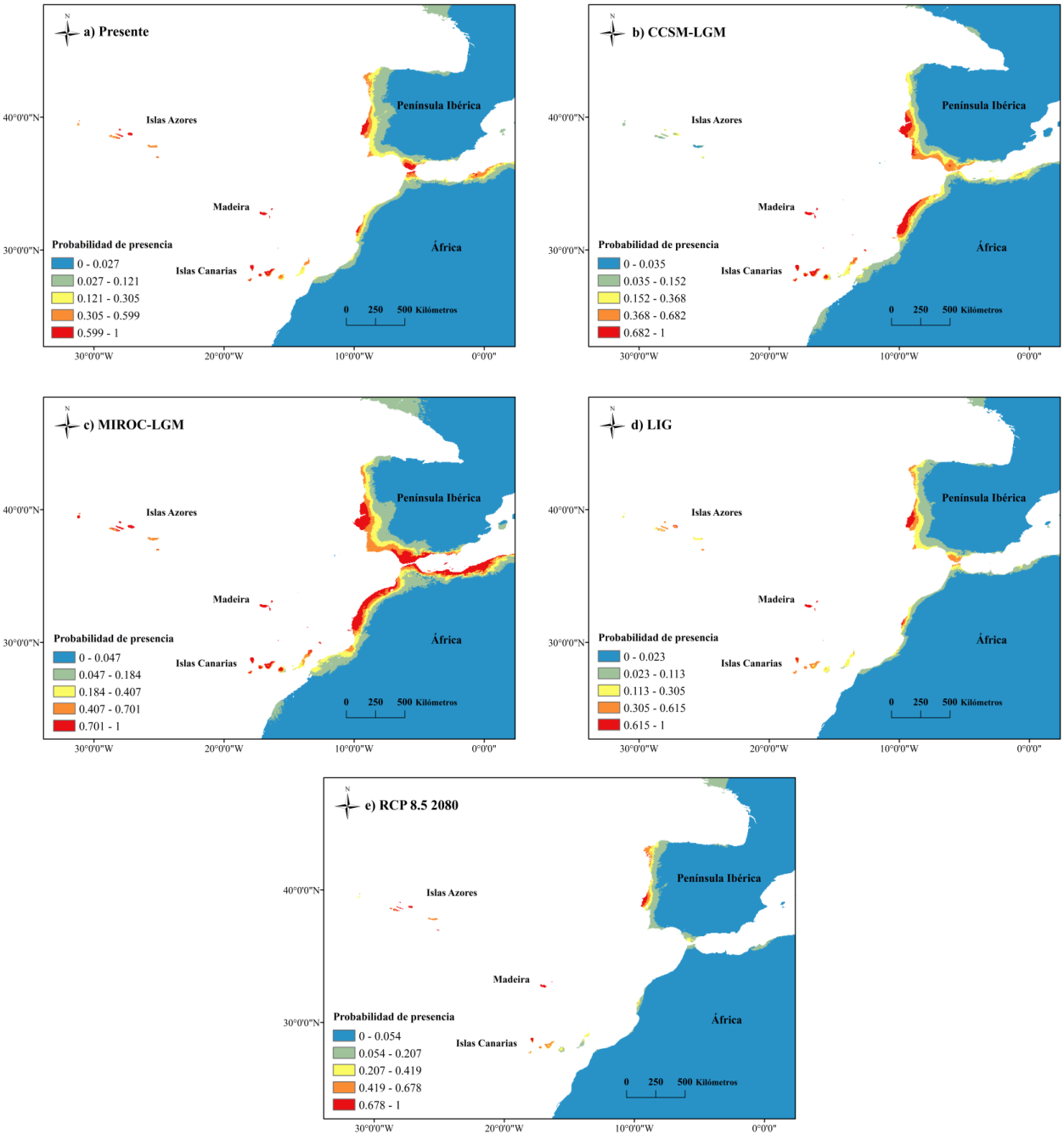


Figure 7. Potential distribution of *D. caudatum* obtained with MAXENT. Top left, at the present time; top right, at the Last Interglacial (LIG, ca. 120,000 years BP); bottom left, at the Last Glacial Maximum (LGM, ca. 21,000 years BP) using the Community Climate System Model (CCSM); bottom right, at 2080 under RCP 8.5 conditions.

4. Discussion

Our results demonstrate the almost absence of asexual reproduction in *D. caudatum* in relation to the species general. However, this is contrasting between regions, as Macaronesian populations show admixed genotypes with supposed sexual reproduction, Andalusia is almost formed by clonal individuals. This is in concordance with the results obtained of genetic diversity, in general high (H_E = 0.495), but very dissimilar between regions, being very low in Andalusian populations.

4.1. Tertiary origin of *D. caudatum*

Recent biogeographical studies of the genus *Diplazium* place its origin in Asia and the beginning of its diversification at 41.7 m.a. These data are consistent with the idea that *Diplazium* species were part of the so-called Paleotropical geoflora (Mai, 1989; 1991; Barrón and Peyrot, 2006). Our dating analysis place *D. caudatum* as part of this paleotropical flora during the Miocene (origin of the lineage 19 m.a.; Figure. 6). The absence of *D. caudatum* in Europe is probably due to the progressive retraction towards the south and subsequent disappearance of the paleotropical geoflora as a result of the strong climatic changes that occurred during the Miocene and Pliocene (Rumsey et al., 2005; Fernández-Palacios et al., 2011; Wei et al., 2015). The Iberian Peninsula is considered the last refuge in Europe of the Paleotropical geoflora, which persisted until the late Pliocene (3.5-2.6 m; Postigo Mijarra et al., 2009), time when the Mediterranean climate was established (3.2 ma; Suc 1984). Of the six responses proposed by Benito Garzón and Sainz de Ollero (2002) of the Paleotropical geoflora to the change towards the Mediterranean climate, *D. caudatum* could be one of the species that survived in Iberian and Macaronesian shelters, like other Tertiary species of ferns (*Culcita macrocarpa, Trichomanes speciosum*, etc). Ben-Menni Schuler et al (in prep) confirmed the Tertiary origin, as *D. caudatum*, of *C. macrocarpa* and *Vandenboschia speciosa* (=*T. speciosum*), and the survival on refuges at the Iberian atlantic coast during glaciations. However, as we will see below, our results support that *D. caudatum* seem to have disappeared from the continent and after been recolonized from Macaronesia.

4.2. Genetic diversity pattern and evolutionary history of *D. caudatum*

The results of our population genetic and phylogeographic analyses suggest an evolutionary history for *D. caudatum* based on the extinction from the European

continent and the survival refuged in Macaronesia, where the species arrived and expanded through long-distance dispersions, and where the diversification of the current gene lineages were due to the isolation between the different archipelagos and probably intra-archipelago as well, and genetic drift. From Macaronesia *D. caudatum* recolonized the European continent.

The first haplotype differentiation event corresponds to the private haplotypes of Madeira (1.8 m.a.). Therefore, we could not affirm the arrival of *D. caudatum* to the Macaronesia before the Pleistocene glaciations (2.5 m.a.) but we could suggest that the lineage differentiation started at the end of the first interglacial. However, we could suppose the arrival before glaciations as this date is consistent with the fossil material from the end of the Pliocene, at the deposit of Sao Jorge in Madeira, of other spermatophyte and fern species of the Tertiary paleoflora that currently occupy the same niches as *D. caudatum* (Heer, 1855). According to Barrón et al. (2003) the arrival of the lauroid forests to the south of the Iberian Peninsula, and with them *D. caudatum*, could occur during the Pliocene, a period where fossils of trees of this paleoflora such as *Laurophyllum* have been recorded and when the region presented available habitats for the persistence of these long-term forests, refugees from the newly established Mediterranean climate (Rodríguez-Sánchez and Arroyo, 2008). Considering the latter, it is possible that the arrival of *D. caudatum* to Macaronesia was during the Pliocene and caused by the climate change that occurred.

Due to the high genetic and haplotypic diversity and the abundance of haplotypes, it is possible that Madeira had a fundamental role in the expansion of *D. caudatum*, acting as a source area for the dispersion to other Macaronesian archipelagos. The demographic Fu's F and Tajima's D analysis supports Madeira, and also Azores, as geographical regions from where *D. caudatum* expands its distribution range. However, BSP supports a general species expansion, but not at any specific region in particular; maybe due to multiple arrivals from the continent and always contributing the same and more frequent haplotypes, therefore these could be more distributed. The subsequent isolation between regions and populations could explain de high number of private alleles and haplotypes. This expansion could be historic rather than recent, since no actual gene flow is perceived.

The absence of a gradient of genetic diversity through Macaronesia discards that colonization has occurred as an unidirectional wave (from more diverse to less diverse areas) and supports the effect of long-distance dispersion, as shown by the not existence of isolation by distance when Andalusian populations where removed from analysis, the fact that F_{ST} values not apparently increase with distance and that the diversification between Azores and the Canary Islands is less than the Canary Islands and Madeira. The geographical distribution of the haplotypes between archipelagos evidences the dispersive capacity of *D. caudatum*, so that not only one haplotype is shared between all the archipelagos, but two and equally distributed, which implies a minimum of two dispersive events between archipelagos. In addition, haplotype HV, very common in the Canary Islands (in IJU and PIJ) and from which the exclusive haplotype of the Canary Islands HIV (in IJU) is derived, is also represented in the Azores Islands, although scarcely (FUR), suggesting a dispersive event at a great distance from the Canary Islands to Azores, and maybe, even recently dispersive events as suggested by BayesAss analysis. F_{ST}, Structure and DAPC analyses support this relation between archipielagos, since according this analyses, Azores and the Canary Islands are genetically very similar. This pattern contrasts with this obtained by Ben-Menni Schuler et al (in prep.) analyzing the phylogeographic history of the Tertiary relict fern *V. speciosa.* According to these authors, Madeira and the Canary Islands are genetically more related, and Azores belongs to a different evolutionary unit.

Recently, the "surfing syngameon hypothesis" (SSH) has been proposed to explain the absence of a diversity gradient in the colonization processes of the macaronesic island systems (Caujapé-Castell, 2011). This hypothesis argues that secondary contacts and subsequent gene flow in island habitats between genotypes that may have previously been isolated on the continent or other island regions generated singameons that enhanced genetic diversity. This hypothesis would explain similar levels of diversity in *D. caudatum* among the three involved archipelagos.

On the other hand, if we exclude haplotypes HII, HIII and HV, the others are private of the different geographical regions (Figure 1.a, Table 3). This evidences a phenomenon of lineage diversification at the archipelagos level, as a consequence of geographical isolation. This is reflected in the plastid and microsatellite AMOVA analysis, in which almost 30% and 10-16% respectively of the genetic variation resides in the interregional component (Table 4). The fact that the divergence of haplotypes coincides with the end

of the Biber-Donau interglacial period and the establishment of the Canary Islands Cold Current (both leading a very arid climate), suggests that the Quaternary glacial cycles, specifically the interglacial periods, as it reflects the MAXENT analysis (Figure 7), could reduce the number of populations and extent of the species, affecting connectivity (decreasing it) between populations within the same archipelago and favoring the isolation and diversification of lineages. The latter would explain that several of the private alleles and haplotypes are exclusive not only at the regional level, but also at the population level within the geographic regions (Table 3).

In relation to the Andalusian populations, the presence of a single haplotype (HI) exclusive to this region, the low genetic diversity levels observed and the genetic-structure results (F_{ST}, DAPC and STRUCTURE analyses) seems to indicate that at some time *D. caudatum* would have disappeared from the Peninsula. The results of the MAXENT analysis show the negative effect of the increase in aridity during interglacial periods in the distribution of *D. caudatum* populations, disappearing from a large number of places. Thus, it is possible that it was during the interglacial periods of the Quaternary when the extinction of the species occured in the European continent. The Sierras of Algeciras is an area recognized as a refuge for the Tertiary paleoflora, also for *D. caudatum* (Benito Garzón and Sainz de Ollero, 2002). However, populations in tertiary shelters have had a long time to have diverged genetically (Rodríguez-Sánchez et al., 2009), so they tend to have a high genetic diversity. The presence of only one haplotype in these populations, their external position in the network, their short time of divergence and the genetic difference in relation to other populations/regions suggest that these populations are the result of a recolonization event of the continent from the Macaronesia. The relation observed between COQ (Andalusia) and the Canary Islands when *K=7* in relation to Azores and Madeira suggests that the Canary Islands could be the source of the Andalusian populations. There are several examples detected by phylogeographic analysis of recolonization of the continent from the Macaronesian archipelagos (Carine et al., 2004; Caujapé-Castell, 2011; Laenen et al., 2010). This Andalusian haplotype has not been detected in the Macaronesian archipelagos maybe because it is has a very low frequency, because it has disappeared, or because it has diverged in situ from HIII and has been established by population isolation and genetic drift in these populations with such a low number of individuals (currently 559 individuals are recognized in Andalusia; Salazar et al., in press).

We conclude that the diversity pattern and phylogeographic history that explains the current distribution of *D. caudatum* along its area is adjusted to a model based on historical long-distance dispersions with gene flow restriction adding recently Azores-Canary Islands connections.

4.3. Implications for conservation

Considering the SDM developed for 2080 under the current greenhouse gas emission conditions (RCP 8.5; Riahi et al., 2011), the area of distribution of the species could be seen reduced, especially in Andalusia and the Canary Islands. *Diplazium caudatum* is a Tertiary relict and therefore the conservation of the species and habitats where it is found is of great importance. Given the isolation observed between geographical regions, evidenced by the AMOVA analysis, the amount of private alleles and haplotypes observed in each of the archipelagos and the homogeneity in microsatellites alleles and the presence of a single private haplotype in Andalusia, we consider that each geographic region could be considered as a different Management Unit (demographically independent populations; Moritz, 1994). Andalusian populations, given their pioneering nature in a phase of recolonization and the low diversity they host deserve special consideration for their protection. In relation to Madeira, this area hosts great diversity, and it would be desirable to give it special consideration for biodiversity conservation and management.

References

Arseneau, J.-R., Steeves, R., Laflamme, M., 2017. Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues. Molecular Ecolology Resources 17, 686–693. https://doi.org/10.1111/1755-0998.12616

Bañares, A., Blanca, G., Güemes, J., Moreno, J. C., Ortiz, S., 2008. Lista roja 2008 de la flora vascular española. Gen. de Medio Natural y Política Forestal (Min. de Medio Ambiente, y Medio Rural y Marino)-SEBICOP, Madrid.

Barrón, E., 2003. Evolución de las floras terciarias en la Península Ibérica 12.

Barrón, E. & Peyrot, D., 2006. La vegetación forestal en el Terciario. In Paleoambientes y cambio climático (pp. 55-76). Fundación Séneca.

Benito Garzón, M. & Sainz de Ollero, H., 2002. Potencialidad del elemento paleotropical lauroide en la Penınsula Ibérica (Doctoral dissertation, MSc Thesis, Universidad Autónoma de Madrid).

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2012. GenBank. Nucleic Acids Research 41, D36–D42. https://doi.org/10.1093/nar/gks1195

Blanca, G., 1999. Libro rojo de la flora silvestre amenazada de Andalucía. Consejería de Medio Ambiente.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Compututational Biology 10, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Carine, M. A., Russell, S. J., Santos-Guerra, A. & Francisco-Ortega, J., 2004. Relationships of the Macaronesian and Mediterranean floras: molecular evidence for multiple colonizations into Macaronesia and back-colonization of the continent in Convolvulus (Convolvulaceae). American Journal of Botany, 91(7), 1070-1085.

Castelloe, J. & Templeton, A. R., 1994. Root probabilities for intraspecific gene trees under neutral coalescent theory. Molecular phylogenetics and evolution, 3(2), 102-113.

Caujapé-Castells, J., 2011. Jesters, red queens, boomerangs and surfers: a molecular outlook on the diversity of the Canarian endemic flora. The biology of island floras, 284-324.

Chanderbali, A. S., van der Werff, H., Renner, S. S., 2001. Phylogeny and historical biogeography of Lauraceae: evidence from the chloroplast and nuclear genomes. Annals of the Missouri Botanical Garden, 104-134.

Clark, P.U., Dyke, A.S., Shakun, J.D., Carlson, A.E., Clark, J., Wohlfarth, B., Mitrovica, J.X., Hostetler, S.W., McCabe, A.M., 2009. The Last Glacial Maximum. Science 325, 710–714. https://doi.org/10.1126/science.1172873

Clement, M., Posada, D. C. K. A., Crandall, K. A., 2000. TCS: a computer program to estimate gene genealogies. Molecular ecology, 9(10), 1657-1659.

Collins, W.D., Bitz, C.M., Blackmon, M.L., Bonan, G.B., Bretherton, C.S., Carton, J.A., Chang, P., Doney, S.C., Hack, J.J., Henderson, T.B., Kiehl, J.T., Large, W.G., McKenna, D.S., Santer, B.D., Smith, R.D., 2006. The Community Climate System Model Version 3 (CCSM3). J. Clim. 19, 2122–2143. https://doi.org/10.1175/JCLI3761.1

Delgado, A. J. & Plaza, L., 2006. Helechos amenazados de Andalucía. Consejería de Medio Ambiente. Junta de Andalucía. Sevilla.

Domínguez–Domínguez, O. & Vázquez, E., 2009. Filogeografía: aplicaciones en taxonomía y conservación. Animal Biodiversity and Conservation 12.

Doyle, J. & Doyle, J. L., 1987. Genomic plant DNA preparation from fresh tissue-CTAB method. Phytochemical Bulletin 19(11), 11-15.

Drummond, A.J., Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Computational Biology 10, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evol. Bioinforma. 1, 117693430500100. https://doi.org/10.1177/117693430500100003

Fernández-Palacios, J.M., de Nascimento, L., Otto, R., Delgado, J.D., García-del-Rey, E., Arévalo, J.R., Whittaker, R.J., 2011. A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests: Palaeo-Macaronesia and the Atlantic island laurel forests. Journal of Biogeography 38, 226–246. https://doi.org/10.1111/j.1365-2699.2010.02427.x

Fick, S.E. & R.J. Hijmans, 2017. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology.

Fu, Y. X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics, 147(2), 915-925.

Gentry-Shields, J., Wang, A., Cory, R.M., Stewart, J.R., 2013. Determination of specific types and relative levels of QPCR inhibitors in environmental water samples using excitation–emission matrix spectroscopy and PARAFAC. Water Resources 47, 3467–3476. https://doi.org/10.1016/j.watres.2013.03.049

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series, 41, 95–98.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22, 160–174. https://doi.org/10.1007/BF02101694

Hasumi, H. & Emori, S., 2004. K-1 Coupled GCM (MIROC) Description, Center for Climate System Research, University of Tokyo.

Heer, O., 1855. Flora tertiaria helvetiae (Vol. 1). J. Wurster.

Hurvich, C. M. & Tsai, C. L., 1993. A corrected Akaike information criterion for vector autoregressive model selection. Journal of time series analysis, 14(3), 271-279.

Jermy, A.C., 1964. Two species of Diplazium in Europe. British Fern Gazette, 9, 160–162.

Koressaar, T. & Remm, M., 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23, 1289–1291. https://doi.org/10.1093/bioinformatics/btm091

Laenen, B., Désamoré, A., Devos, N., Shaw, A. J., González-Mancebo, J. M., Carine, M. A., & Vanderpoorten, A., 2011. Macaronesia: a source of hidden genetic diversity for post-glacial recolonization of western Europe in the leafy liverwort Radula lindenbergiana. Journal of Biogeography, 38(4), 631-639.

Mai, D. H., 1989. Development and regional differentiation of the European vegetation during the Tertiary. In Woody plants—evolution and distribution since the Tertiary (pp. 79-91). Springer, Vienna.

Mai, D. H., 1991. Palaeofloristic changes in Europe and the confirmation of the Arctotertiary-Palaeotropical geofloral concept. Review of Palaeobotany and Palynology, 68(1-2), 29-36.

Moritz, C., 1994. Defining 'evolutionarily significant units' for conservation. Trends in ecology & evolution, 9(10), 373-375.

Morris, A.B. & Shaw, J., 2018. Markers in time and space: A review of the last decade of plant phylogeographic approaches. Molecular Ecology 27, 2317–2333. https://doi.org/10.1111/mec.14695

Moya, P., Delgado, A., Plaza, L., 2010. Propagación y aclimatación de cuatro especies de Pteridófitos amenazados en la provincia de Cádiz 13.

Nei, M., 1987. Molecular evolutionary genetics. Columbia university press. New York

Otto-Bliesner, B.L., 2006. Simulating Arctic Climate Warmth and Icefield Retreat in the Last Interglaciation. Science 311, 1751–1753. https://doi.org/10.1126/science.1120808

Petit, R. J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D., Vendramin, G. G., 2005. Invited review: comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Molecular Ecology, 14(3), 689-701.

Petit, R. J. & Hampe, A., 2006. Some evolutionary consequences of being a tree. Annual Review of Ecology, Evolution, and Systematics., 37, 187-214.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecological Modellling 190, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Posada, D., 2008. jModelTest: Phylogenetic Model Averaging. Molecular Biology and Evolution 25, 1253–1256. https://doi.org/10.1093/molbev/msn083

Postigo Mijarra, J.M., Barrón, E., Gómez Manzaneque, F., Morla, C., 2009. Floristic changes in the Iberian Peninsula and Balearic Islands (south-west Europe) during the Cenozoic. Journal of Biogeography 36, 2025–2043. https://doi.org/10.1111/j.1365-2699.2009.02142.x

Pulido, F., Valladares, F., Calleja, J.A., Moreno, G., González-Bornay, G., 2008. Tertiary relict trees in a Mediterranean climate: abiotic constraints on the persistence of *Prunus lusitanica* at the eroding edge of its range. Journal of Biogeography 35, 1425–1435. https://doi.org/10.1111/j.1365-2699.2008.01898.x

Rambaut, A. 2016. FigTree v1. 4.3 software. Institute of Evolutionary Biology, University of Edinburgh.

Rambaut A., Drummond A.J., Xie D., Baele G. and Suchard M.A., 2018 Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. Systematic Biology. syy032. doi:10.1093/sysbio/syy03

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., Rafaj, P., 2011. RCP 8.5—A scenario of comparatively high greenhouse gas emissions. Climate Change 109, 33–57. https://doi.org/10.1007/s10584-011-0149-y

Rodríguez-Sánchez, F.& Arroyo, J., 2008. Reconstructing the demise of Tethyan plants: climate-driven range dynamics of *Laurus* since the Pliocene. Global Ecology and Biogeography 17, 685–695. https://doi.org/10.1111/j.1466-8238.2008.00410.x

Rodríguez-Sánchez, F., Guzmán, B., Valido, A., Vargas, P., Arroyo, J., 2009. Late Neogene history of the laurel tree (Laurus L., Lauraceae) based on phylogeographical analyses of Mediterranean and Macaronesian populations. Journal of Biogeography, 36(7), 1270-1281.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Systematic Biology 61, 539–542. https://doi.org/10.1093/sysbio/sys029

Rothfels, C.J., Sundue, M.A., Kuo, L., Larsson, A., Kato, M., Schuettpelz, E., Pryer, K.M., 2012. A revised family-level classification for eupolypod II ferns (Polypodiidae: Polypodiales). Taxon, 61, 515–533.

Rumsey, F.J., Barrett, J.A., Gibby, M., Russell, S.J., Vogel, J.C., 2005. Reproductive strategies and population structure in the endangered pteridophyte *Trichomanes speciosum* (Hymenophyllaceae: Pteridophyta) 11.

Salazar, C., Delgado, A.,Marisca, D, Pérez Latorre, A.V.,Plaza, L., Rendón, J.L., Suárez-Santiago, V.N., Sánchez, T. y Mesa, R., 2017. Diplazium caudatum (Cav.) Jermy. en Bañares, Á., G. Blanca, J. Güemes, J.C. Moreno & S. Ortiz, eds. Atlas y Libro Rojo de la Flora Vascular Amenazada de España. Adenda 2017. Dirección General de Medio Natural y Política Forestal (Ministerio de Medio Ambiente, y Medio Rural y Marino)-Sociedad Española de Biología de la Conservación de Plantas. Madrid. En prensa

Shaw, J., Lickey, E.B., Schilling, E.E., Small, R.L., 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. American Journal of Botany 94, 275–288. https://doi.org/10.3732/ajb.94.3.275

Shaw, J., Small, R.L., Lickey, E.B., Hauk, W.D., 2005. Amplification of noncoding chloroplast DNA for phylogenetic studies in lycophytes and monilophytes with a comparative example of relative phylogenetic utility from Ophioglossaceae. Molecular Phylogenetics and Evolution 36, 509–522. https://doi.org/10.1016/j.ympev.2005.04.018

Smith, S. A., & Donoghue, M. J., 2008. Rates of molecular evolution are linked to life history in flowering plants. Science, 322(5898), 86-89.

Suc, J. P., 1984. Origin and evolution of the Mediterranean vegetation and climate in Europe. Nature, 307(5950), 429.

Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Molecular Biology 17, 1105–1109. https://doi.org/10.1007/BF00037152

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123(3), 585-595.

Templeton A.R., Crandall K.A., Sing C.F., 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132, 619–633.

Turner, C., 2002. Formal Status and Vegetational Development of the Eemian Interglacial in Northwestern and Southern Europe. Quaternary Research 58, 41–44. https://doi.org/10.1006/qres.2002.2365

Tryon, R.M. & Tryon, A.F., 1982. Ferns and allied plants with special reference to tropical America. Springer, New York.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3—new capabilities and interfaces. Nucleic Acids Research 40, e115–e115. https://doi.org/10.1093/nar/gks596

Waples, R. S., 1991. Pacific salmon, *Oncorhynchus spp*., and the definition of" species" under the Endangered Species Act. Marine Fisheries Review, 53(3), 11-22.

Wei, R., Schneider, H., Zhang, X.-C., 2013. Toward a new circumscription of the twinsorus-fern genus *Diplazium* (Athyriaceae): A molecular phylogeny with morphological implications and infrageneric taxonomy. Taxon 62, 441–457. https://doi.org/10.12705/623.8

Wei, R., Xiang, Q., Schneider, H., Sundue, M.A., Kessler, M., Kamau, P.W., Hidayat, A., Zhang, X., 2015. Eurasian origin, boreotropical migration and transoceanic dispersal in the pantropical fern genus *Diplazium* (Athyriaceae). Journal of Biogeography 42, 1809–1819. https://doi.org/10.1111/jbi.12551

Wei, R., Yan, Y.-H., Harris, A., Kang, J.-S., Shen, H., Xiang, Q.-P., Zhang, X.-C., 2017. Plastid Phylogenomics Resolve Deep Relationships among Eupolypod II Ferns with Rapid Radiation and Rate Heterogeneity. Genome Biology and Evolution 9, 1646–1657. https://doi.org/10.1093/gbe/evx107

Xin, Z.& Chen, J., 2012. A high throughput DNA extraction method with high yield and quality. Plant Methods 8, 26. https://doi.org/10.1186/1746-4811-8-26

Xu, B.& Yang, Z., 2013. PAMLX: A Graphical User Interface for PAML. Molecular Biology and Evolution 30, 2723–2724. https://doi.org/10.1093/molbev/mst179

Yang, Z., 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. Molecular Biology and Evolution 24, 1586–1591. https://doi.org/10.1093/molbev/msm088

Supplementary data

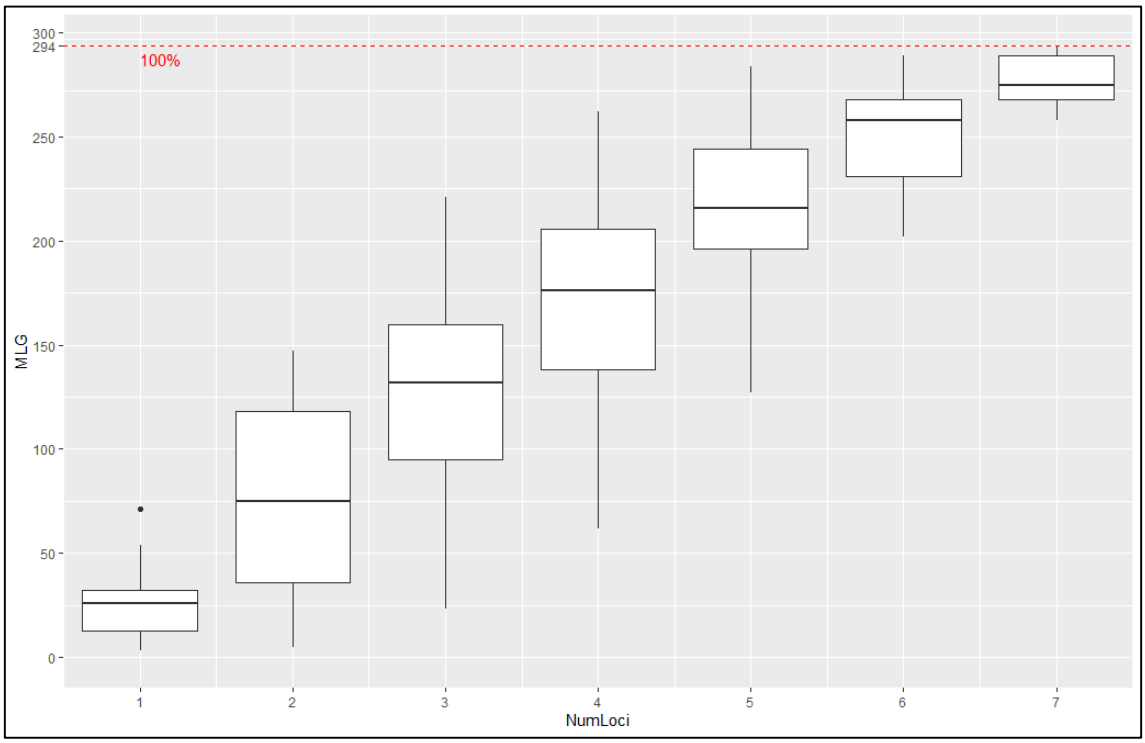


Figure S1. Genotypic accumulation curve showing the resolutive power of the eight microsatellite used in this study.

Table S1. Percent contribution and permutation importance (MaxEnt) of selected model. Variables in bold were selected for the final model.

Variable	MaxEnt Percent contribution	MaxEnt Permutation importance
Mean Diurnal Range	**52.5**	**26.2**
Min. Temperature of Coldest Month	**24.1**	**70.3**
Precipitation of Warmest Quarter	16.7	3.2
Type of Soil	2.5	0.1
Precipitation of Driest Month	1.9	0.0
Precipitation Seasonality	1.8	0.0
Precipitation of Coldest Quarter	0.3	0.1
Max. Temperature of Warmest Month	0.2	0.1
Precipitation of Wettest Month	0.1	0.0

Table S2. Pairwise population F_{ST} for microsatellites. Values in bold were significant at the 5% nominal level after sequential Bonferroni correction.

	CRM	COQ	AGU	CID	FUR	SER	CED	PIJ	IJU	ANC	FRI	POR
CRM	--	0.306	0.504	0.544	0.540	0.503	0.566	0.481	0.566	0.506	0.317	0.310
COQ	**0.586**	--	**0.295**	**0.328**	**0.338**	**0.283**	**0.373**	**0.304**	**0.350**	**0.378**	**0.284**	**0.314**
AGU	**0.631**	**0.315**	--	**0.086**	**0.075**	**0.081**	**0.111**	**0.058**	**0.163**	**0.263**	**0.166**	**0.224**
CID	**0.667**	**0.353**	**0.089**	--	**0.056**	**0.072**	**0.123**	**0.110**	**0.194**	**0.292**	**0.183**	**0.237**
FUR	**0.675**	**0.364**	**0.071**	**0.057**	--	0.049	**0.099**	**0.082**	**0.188**	**0.273**	**0.176**	**0.228**
SER	**0.653**	**0.314**	**0.088**	**0.082**	0.058	--	**0.121**	**0.123**	**0.197**	**0.225**	**0.199**	**0.246**
CED	**0.674**	**0.385**	**0.102**	**0.126**	**0.098**	**0.123**	--	**0.142**	**0.180**	**0.336**	**0.178**	**0.244**
PIJ	**0.616**	**0.325**	**0.057**	**0.108**	**0.075**	**0.131**	**0.138**	--	**0.170**	**0.247**	**0.156**	**0.199**
IJU	**0.685**	**0.374**	**0.169**	**0.209**	**0.204**	**0.220**	**0.178**	**0.180**	--	**0.330**	**0.175**	**0.223**
ANC	**0.635**	**0.399**	**0.268**	**0.301**	**0.279**	**0.240**	**0.334**	**0.254**	**0.341**	--	**0.266**	**0.288**
FRI	**0.478**	**0.317**	**0.166**	**0.195**	**0.192**	**0.222**	**0.175**	**0.160**	**0.179**	**0.270**	--	**0.027**
POR	**0.477**	**0.349**	**0.225**	**0.251**	**0.248**	**0.273**	**0.244**	**0.206**	**0.228**	**0.287**	**0.029**	--

Table S3. Mean recent migration rates (m) among the studied populations. estimated from eight microsatelliteloci using the BAYESASS program. Values on the diagonal (underlined) indicate the proportion of individuals in each generation that are not migrants. Values in bold are the *m* rates that are informative.

	From											
To	CRM	COG	AGU	CID	FUR	SER	CED	PIJ	IJU	ANC	FRI	POR
CRM	0.689 (0.021)	0.085 (0.037)	0.023 (0.021)	0.021 (0.020)	0.022 (0.021)	0.022 (0.020)	0.024 (0.024)	0.021 (0.019)	0.022 (0.021)	0.022 (0.021)	0.022 (0.021)	0.022 (0.02)
COQ	0.016 (0.016)	0.791 (0.049)	0.037 (0.027)	0.015 (0.014)	0.016 (0.016)	0.014 (0.014)	0.028 (0.038)	0.016 (0.015)	0.015 (0.015)	0.016 (0.016)	0.016 (0.015)	0.015 (0.014)
AGU	0.008 (0.007)	0.008 (0.008)	0.909 (0.023)	0.008 (0.008)	0.007 (0.007)	0.008 (0.007)	0.010 (0.010)	0.007 (0.007)	0.008 (0.008)	0.008 (0.007)	0.008 (0.008)	0.008 (0.007)
CID	0.008 (0.008)	0.008 (0.008)	**0.243** (0.022)	0.674 (0.008)	0.008 (0.008)	0.008 (0.007)	0.008 (0.007)	0.007 (0.007)	0.008 (0.008)	0.008 (0.007)	0.008 (0.007)	0.008 (0.007)
FUR	0.009 (0.008)	0.009 (0.009)	**0.224** (0.025)	0.008 (0.008)	0.675 (0.008)	0.008 (0.008)	0.013 (0.012)	0.009 (0.008)	0.008 (0.008)	0.015 (0.012)	0.008 (0.008)	0.008 (0.008)
SER	0.009 (0.009)	0.009 (0.008)	**0.236** (0.025)	0.008 (0.008)	0.008 (0.008)	0.675 (0.008)	0.010 (0.010)	0.008 (0.008)	0.008 (0.008)	0.009 (0.009)	0.008 (0.008)	0.008 (0.008)
CED	0.009 (0.009)	0.009 (0.008)	0.079 (0.033)	0.008 (0.007)	0.008 (0.008)	0.008 (0.008)	0.835 (0.035)	0.008 (0.008)	0.008 (0.008)	0.008 (0.008)	0.008 (0.008)	0.008 (0.008)
PIJ	0.008 (0.008)	0.008 (0.008)	**0.242** (0.026)	0.007 (0.007)	0.007 (0.007)	0.007 (0.007)	0.008 (0.008)	0.677 (0.015)	0.009 (0.009)	0.008 (0.008)	0.007 (0.007)	0.008 (0.007)
IJU	0.008 (0.008)	0.008 (0.008)	**0.212** (0.057)	0.008 (0.008)	0.008 (0.008)	0.008 (0.008)	0.009 (0.008)	0.008 (0.008)	0.704 (0.052)	0.008 (0.008)	0.008 (0.008)	0.008 (0.007)
ANC	0.008 (0.007)	0.008 (0.008)	0.027 (0.015)	0.007 (0.007)	0.008 (0.008)	0.008 (0.008)	0.009 (0.008)	0.007 (0.007)	0.008 (0.008)	0.892 (0.024)	0.007 (0.007)	0.008 (0.008)
FRI	0.008 (0.008)	0.013 (0.012)	0.02 (0.015)	0.009 (0.009)	0.009 (0.009)	0.009 (0.009)	0.024 (0.019)	0.009 (0.009)	0.019 (0.020)	0.009 (0.009)	0.844 (0.040)	0.022 (0.021)
POR	0.008 (0.008)	0.01 (0.009)	0.01 (0.01)	0.008 (0.007)	0.008 (0.007)	0.008 (0.008)	0.010 (0.009)	0.009 (0.009)	0.012 (0.014)	0.009 (0.008)	0.041 (0.025)	0.864 (0.032)

Table S4. Neutrality tests Fu's F and Tajima's D of the populations and regional groups of *D. caudatum* performed to measure population expansion and contraction. Values in bold are significant; N.A.: not applicable

Region/Population	Fu's F		Tajima's D	
	F	**P-value**	***D***	**P-value**
Andalusia	0.000	N.A.	0.000	1
CRM	0.000	N.A.	0.000	1
COQ	0.000	N.A.	0.000	1
Azores	-1.625	0.093	-0.393	0.383
AGU	0.626	0.511	1.225	0.937
CID	0.626	0.526	1.225	0.935
FUR	**-1.938**	0.014	-0.972	0.1
SER	-0.474	0.200	0.243	0.769
Canary Islands	0.039	0.514	0.554	0.744
ANC	0.000	N.A.	0.000	1
CED	0.626	0.513	1.225	0.930
PIJ	1.687	0.761	1.459	0.952
IJU	**-1.648**	0.047	-0.175	0.439
Madeira	-1.587	0.057	-0.279	0.381
FRI	0.061	0.320	-0.175	0.432
POR	**-1.648**	0.047	-0.175	0.432

Chapter 5. Development of polymorphic microsatellite markers for the Iberian-Macaronesian endemic Laurisila Brake (*Pteris incompleta*, Pteridaceae).

Samira Ben-Menni Schuler, Ana Teresa Romero García, Víctor N. Suárez Santiago

Manuscript in preparation

Abstract

Premise of the study:

We characterize 8 microsatellite loci in the fern *Pteris incompleta* (Pteridaceae), enabling studies on the genetic population structure of this Iberian-Macaronesian species using DNA hypervariable markers.

Methods and Results:

Eight primer sets were developed and tested on 47 individuals in a total of one Azorean and one Madeiran population of *P. incompleta*. The primers amplified di, tri- and tetranucleotide repeats. The number of alleles per locus ranged from four to twenty-two (average: 13.375), and the expected heterozygosity ranged from 0.380 to 0.943 among the populations analyzed.

Conclusions:

The 8 microsatellite markers developed will be useful in characterizing the genetic diversity of *P. incompleta* and understanding its population structure and biogeographic history, and will provide important genetic data for the conservation of this species.

Pteris incompleta Cav. (Pteridaceae) is an endemic Iberian-Macaronesian fern. Its populations are restricted to the Macaronesian Islands and the southwest of the Iberian Peninsula (Cádiz, Andalusia, Spain), in places considered shelters during the Quaternary glacial cycles (Benett et al., 1991; Fernández-Palacios et al., 2010). It lives in shady and humid places associated to permanent courses of water with abundant riparian vegetation. No previous studies on the population genetics of *P. incompleta* have been done, therefore, no information about the genetic diversity status is avaliable, even for the critically endangered populations of the Iberian Peninsula.

The distribution and the history of *P. incompleta* make this fern an interesting plant in population genetics and biogeographic studies. The aim of this paper was to develop a suite of microsatellite markers for *P. incompleta* that will allow a detailed analysis of its genetic diversity and structuring throughout its distribution. Eight microsatellite

markers were developed following the microsatellite-enrichment procedure based on Reddy et al. (2001).

Methods and Results

Genomic DNA was extracted from fresh fronds of *P. incompleta* from one sample of the population Ribeiro Frio (FRI), Madeira, Portugal (voucher deposited at the herbarium of the Universidad de Granada [GDA]: GDA XX; georeference: N32.734°/W16.886°) using the NucleoSpin Plant Kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany). We digested 2 µg of genomic DNA in a single reaction mixture with *Hae*III and *Rsa*I (20 U of each), and ligated AP11 and AP12 adapters (100 ng of each). The ligation product (2 µL) was amplified by polymerase chain reaction (PCR) using 100 ng of the single primer AP11, 2 mm $MgCl_2$, 1 mm dNTPs and 3 U of *Taq* DNA polymerase (Promega). The resulting fragment library was enriched by hybridizing 5 µL of DNA amplified fragments to 5′-biotin-labelled oligonucleotide probes $(GA)_{20}$ and $(CT)_{30}$. Probe-target fragments were captured using streptavidin-coated magnetic beads (Promega), and then purified. Five microlitres of purified DNA sample was used as a template for PCR using primer AP11 as described above. Four microlitres of the PCR was ligated into a TOPO TA PCR vector (Invitrogen), cloned into One Shot *Escherichia coli* competent cells (Invitrogen) and selected by growing them overnight on kanamycin LB-agar plates. Selected colonies were screened for recombinant plasmid by PCR using M13 forward and reverse primers, and cultured in LB overnight for subsequent sequencing using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) in an ABI 3100-Avant Genetic Analyser (Applied Biosystems). Nucleotide sequences were edited with the SeqMan II version 3.61 program from the DNAStar software package (LASERGEN).

A total of 177 colonies were sequenced, of which 105 contained microsatellites. For validation tests, 40 PCR primer pairs were designed, using the Primer3 program (Rozen and Skaletsky, 2000) or manually, and used for amplification in eight DNA samples of *P. incompleta* from two populations: FRI and Algar do Carvão (CAR), Terceira, Azores, Portugal (voucher deposited at GDA: GDA XX; georeference: N38.727°/W27.215°). Forward primers were fluorescently labelled in their 5′ ends (6FAM, HEX, ATTO-550). Forward primers were synthesized with M13 tails (5′-

CACGACGTTGTAAAACGAC-3′) preceding the 5′ end to facilitate cost-efficient fluorescent labelling of PCR products. DNA was isolated as above and PCR optimizations were performed by modifying the range of annealing temperatures. PCR reactions were performed in 10-μL reactions containing 20 ng of genomic DNA, 0.02 μM of the M13-labeled forward primer, 0.45 μM of each reverse primer and fluorolabeled M13 primer (5′-6FAM, 5′-HEX, 5′-ATTO-550), 5 μL Kapa 2G Robust HotStart ReadyMix (Kapa Biosystems, Boston, MA). Cycling parameters consisted of 3 min of denaturing at 94°C; followed by 35 cycles of 94°C for 30 s, 50−58°C for 30 s, and 72°C for 30 s; and a final extension at 72°C for 10 min. PCR products were analyzed with GeneScan 500 LIZ Size Standard (Applied Biosystems, Foster City, California, USA) on an ABI 3100-Avant Genetic Analyzer (Applied Biosystems) at Genoscreen (Lille, France). Microsatellite data were analyzed with GeneMarker version 1.51 (SoftGenetics, State College, Pennsylvania, USA). Of the 40 microsatellite markers tested, 25 gave a positive signal on electrophoretic gels; of these, 9 yielded an ambiguous allelic pattern, 8 were monomorphic, and 8 were polymorphic with easily interpretable electropherograms. The characteristics of the 8 polymorphic loci are shown in Table 1.

Table 1. Characteristics of 8 microsatellite loci developed in *Pteris incompleta.*

Locus	Primer sequences (5′–3′)[a]	Repeat motif	Allele size range (bp)	T_a (°C)	EMBL accession no.
PI-95	F: ATAGCGGGCTCAGAGAGTTC		192-232	58	
	R: GTCATCTATACCCACGAAAT	(TG)15			
PI-CA2	F: GTATAGCTAGTTAGAC		172-178	50	
	R: GGTTGCAAAGTAAACACTTC	(TATG)4 +(GT)14			
PI-AT13	F: GTACCTTGCAGAAGCCGC	(CA)2(AGA)2A(AG)10	178-188	58	
	R: AGCTATGGTGTTTTGCTTTT				
PI-AT30	F: CATGGAGGGGTGGGTCAG		184-214	58	
	R: CTACCCCACGCGTTTTCTTC	(GA)12			
PI-AT51	F: GATACGCAGCAACCCTTCAG		249-297	58	
	R: AGAGCTTTCCGGCCATTTTC	(GA)19			
PI-AT32m1	F: ACACGGATACAGTTGGCCAT		199-243	58	
	R: AAATGCAAGCGACACCATGT	(GA)14			
PI–AT36	F: GTGTATGTGCGTTTGTAAGTGC		228-264	58	
	R: GGTGGAGCTAGCATAACCTACA	(TG)17			
PI-CA46	F: GCTAAGCAAATGTCACAGAA		208-262	55	
	R: GGTATAGATAGCAAGGATTGTC	(CA)4TA(CA)16			

Note: EMBL = European Molecular Biology Laboratory; T_a = annealing temperature.

[a] Forward primer sequence and allele size range do not include the M13-tail sequence.

Genetic diversity of the selected loci was assessed by genotyping (as above) a total of 47 individuals of the two aforementioned populations (17 from CAR, 30 from FRI). Data were analyzed with the software GenoDive version 2.0b24 (Meirmans and Van Tienderen, 2004). In total, 46 unique multilocus genotypes (MLG) were identified (CAR: 17, FRI: 29) by using the Assign Clones function and the Stepwise Mutation Model for calculating the interindividual distances in GenoDive. The average number of alleles per locus was 13.375 (CAR: 7.125, FRI: 11.875). No locus was monomorphic in any population. Considering the 46 MLG, the expected heterozygosity across all the populations analyzed was 0.791 (CAR: 0.724, FRI: 0.805; intrapopulation

heterozygosity: 0.765). The values for allele number and heterozygosity for each locus and population are shown in Table 2.

Table 2. Summary genetic statistics for the two populations of *Pteris incompleta* analyzed with the 8 newly developed microsatellites.[a]

	CAR (n = 17)			FRI (n = 29)		
Locus	A	H_o	H_e	A	H_o	H_e
PI-95	8	0.353	0.757	17	0.867	0.909
PI-CA2	4	0.471	0.443	3	0.533	0.635
PI-AT13	3	0.294	0.437	6	0.367	0.38
PI-AT30	7	0.353	0.855	14	0.867	0.912
PI-AT51	12	0.882	0.904	19	0.867	0.943
PI-AT32m1	7	0.294	0.881	13	0.467	0.882
PI–AT36	7	0.471	0.673	14	0.7	0.905
PI-CA46	9	0.059	0.842	9	0.2	0.872

Note: A = number of alleles sampled; H_e = expected heterozygosity; H_o = observed heterozygosity; n = number of individuals (genets) sampled.

a Locality and voucher information for the sampled populations: Algar do Carvão, Terceira, Azores, Portugal (voucher: GDA XX; georeference: N38.727°/W27.215°); Ribeiro Frio, Madeira, Portugal (voucher: GDA XX; georeference: N32.734°/W16.886°).

Conclusions

The 8 microsatellites described in this study will enable population genetics studies of *P. incompleta* throughout its distribution. These microsatellites will be used to determine the genetic diversity within and among the different populations of *P. incompleta*. This information will be essential for the correct management of this endangered species. Moreover, future studies using these 8 microsatellite markers will allow questions about the biogeographic history of *P. incompleta* to be tested, such as the refugial status of its populations. Our results in this study show two populations—considered glacial shelters (Fernández-Palacios et al., 2010)—to be moderately-highly

diverse. The critically endangered status of *P. incompleta* in the southwest of the Iberian Peninsula, the only place in continental Europe where this species exist, makes necessary a genetic analysis of these populations to be able to take protective measures if necessary.

References

Bennett, K. D., Tzedakis, P. C., and Willis, K. J. 1991. Quaternary refugia of north European trees. *Journal of Biogeography* 18: 103-115.

Fernández-Palacios, J. M., de Nascimento, L., Otto, R., Delgado, J. D., García-del-Rey, E., Arévalo, J. R., and Whittaker, R. J. 2011. A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography* 38: 226-246.

Meirmans, P. G., and Van Tienderen, P. H. 2004. GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4: 792-794.

Reddy, O. U. K., Pepper, A. E., Abdurakhmonov, I., Saha, S., Jenkins, J. N., Brooks, T., Bolek, Y., and El-Zik, K. M. 2001. New dinucleotide and trinucleotide microsatellite marker resources for cotton genome research. *Journal of Cotton Sciences* 5: 103-113.

Rozen, S., and Skaletsky, H. J. 2000. Primer3 on the WWW for general users and for biologist programmers. In S. Misener and S. Krawetz [eds.], Bioinformatics Methods and Protocols, 365-386. Humana Press, Totowa, New Jersey, USA.

Discusión

A lo largo de esta tesis hemos analizado la diversidad genética y el patrón filogeográfico de cuatro helechos probablemente relictos terciarios, con mayor o menor grado de amenaza, incluyendo: a) probar la hipótesis de un posible origen terciario en *Vandenboschia speciosa*, *Culcita macrocarpa* y *Diplazium caudatum*, b) determinar la diversidad genética y el patrón filogeográfico a lo largo de todo su rango de distribución en *V. speciosa*, *C. macrocarpa* y *D. caudatum* y explorar cómo los cambios climáticos en el pasado pudieron afectar al modelado de esos patrones; en *Pteris incompleta* comparamos la diversidad genética en dos poblaciones macaronésicas, y c) determinar el posible impacto del calentamiento climático futuro sobre la diversidad genética y la distribución de especies en *V. speciosa*, *C. macrocarpa* y *D. caudatum*. A continuación, resumimos los principales hallazgos y discutimos las conclusiones de los Capítulos 1 a 5.

En esta tesis, determinamos y confirmamos el origen terciario de *V. speciosa*, *C. macrocarpa* y *D. caudatum*, posiblemente durante el Mioceno. Estos tres helechos podrían formar parte de la llamada Geoflora Paleotropical (Mai, 1989; 1991; Barrón & Peyrot, 2006), una flora dominada por plantas perennes y termófilas que poblaron el hemisferio norte desde el Cretácico tardío hasta el Mioceno tardío (Mai, 1989, 1991), siendo los helechos el componente principal de la capa herbácea (Pichi-Sermolli, 1979, 1991; Barrón, 2003; Barrón y Peyrot, 2006). Durante el Plioceno-Pleistoceno, la zona donde habitaba la Geoflora Paleotropical sufrió un cambio drástico en el clima volviéndose más seco y frío. La presencia o más bien supervivencia actual de elementos típicos de bosques de lauoroides en Europa, como son las especies de helechos estudiados en esta tesis, puede ser debido a la existencia de refugios microclimáticos con condiciones adecuadas de temperatura y humedad distribuidos a lo largo de la costa atlántica europea, sobre todo de la Pensínsula Ibérica y las islas Macaronésicas: Azores, Islas Canarias y Madeira, donde se encuentran actualmente.

V. speciosa es la única especie de las cuatro estudiadas que posee ambas generaciones de su ciclo de vida (gametofito y esporofito) perennes, y es la que presenta un rango de distribución más amplio; sin embargo la distribución geográfica de ambas generaciones

no es similar. En esta tesis se ha comprobado que a pesar de esta diferente distribución gametofito-esporofito, no existe diferenciación o estructuración genética entre ambas fases, al igual que señaló Rumsey et al. (1999).

En esta tesis se proponen diferentes escenarios para explicar los patrones actuales de diversidad, distribución y filogeográficos de las especies estudiadas. En las tres especies que se analizó su patrón filogeográfico, parece ser que los eventos geológico-climáticos acaecidos durante el Mioceno-Plioceno y posteriormente el Pleistoceno son las causantes de su actual diversidad genética y distribución aunque las respuestas por parte de cada especie son diferentes. Uno de los patrones más interesantes es el obtenido para *V. speciosa*, según el cual la especie muestra la presencia de dos unidades evolutivas principales con una marcada diferenciación norte-sur, tal y como sugirieron Rumsey et al., 1996. En relación a las poblaciones continentales de *C. macrocarpa* y las localizadas en la costa atlántica europea de *V. speciosa* parecen haber sobrevivido a estos eventos adversos, aunque en lugares pequeños y climáticamente favorables bajo el fuerte efecto de la deriva genética como muestra nuestro estudio de la diversidad genética de *V. speciosa* en las poblaciones esporofíticas de Cádiz, Andalucía (Ben-Menni Schuler et al., 2017). Sin embargo, *D. caudatum* parece haber desaparecido de la única región en la que se encuentra actualmente, Cádiz, (Andalucía) y su presencia actual se podría deber a una recolonización reciente de poblaciones procedentes de la Macaronesia. En todos los casos, las poblaciones de distintas regiones parecen estar aisladas unas de otras, aunque ese aislamiento parece más evidente en *D. caudatum*, dónde las poblaciones andaluzas son completamente diferentes a las restantes y las maderenses también muestran una diferencia significativa.

Parece claro, que una de las regiones refugio más importantes durante estos periodos de frio y aridez que caracterizan el final del Mioceno y las posteriores glaciaciones es toda la Cornisa Cantábrica, ya que las poblaciones más diversas en *V. speciosa* y *C. macrocarpa* se encuentran en esa región. Esto es corroborado en otras especies incluyendo helechos (Jermy, 1984; Ramil-Rego & Gómez-Orellana, 1996; Vogel et al., 1999; Hunt et al., 2009; Beatty & Provan, 2013, 2014). También parece probable que en *V. speciosa*, durante los periodos interglaciares, la cornisa cantábrica haya servido de fuente para re/colonizar poblaciones situadas más al norte y el este que pudieran haber desaparecido en condiciones climáticas más adversas.

Los resultados obtenidos en esta tesis muestran que las poblaciones macaronésicas de estos helechos sobrevivieron a las glaciaciones del Pleistoceno y por lo tanto se consideran refugio. Nuestros resultados sugieren que los archipiélagos macaronésicos fueron colonizados durante el Plio-Pleistoceno, como está confirmado por otras angiospermas relictas (Kondraskov et al., 2015). Respecto a las relaciones existentes entre las poblaciones macaronésicas y las continentales, resulta interesante comprobar que en *V. speciosa*, las poblaciones de Azores se relacionan con la unidad evolutiva del norte, mientras que las Islas Canarias y Madeira se encuentran más relacionadas a la unidad evolutiva del sur, patrón también observado en *Asplenium hemionitis* (Duraes, 2005). Sin embargo, en *C. macrocarpa* las poblaciones de Azores y las Islas Canarias se encuentran relacionadas y éstas con las de Andalucía, presentándose una mayor divergencia con las poblaciones de la Cornisa Cantábrica. *D. caudatum* también presenta relacionadas las poblaciones de Azores y las Islas Canarias, sin embargo, las poblaciones andaluzas y maderenses muestran una clara diferenciación del resto. Estos diferentes patrones de diversidad y estructuración pueden ser debidos a las distintas poblaciones que han servido de fuente para colonizar los distintos archipiélagos macaronésicos. Además, pueden haber influido las conexiones submarinas existentes entre los archipielagos, y éstos y el continente que durante las glaciaciones y su salida a la superficie por la bajada del nivel del mar pudieron haber servido de puentes para favorecer el flujo génico.

Se ha observado en *P. incompleta* una elevada diversidad genética en las dos poblaciones analizadas, ambas pertenecientes a la macaronesia. Es de destacar la elevada diversidad genética que presentan las poblaciones macaronésicas de todos los helechos analizados, remarcando el carácter relictual de esta región.

En relación a las perspectivas de futuro que presentan estas especies en general y sus poblaciones en particular, es de destacar que un aumento de las temperaturas y un escenario de tipo RCP 8.5 según el mostrado por el IPCC (Van Vuuren et al., 2011) afectaría considerablemente la distribución y diversidad de éstas. Es común en todas el desplazamiento hacia el norte de la disponibilidad de áreas adecuadas para su presencia, pudiendo desaparecer las poblaciones del sur de la Península Ibérica y las Islas Canarias. Las especies analizadas en esta tesis se han incluido en listas rojas con

diferentes niveles de protección; sin embargo, recientemente se han realizado estudios exhaustivos sobre la demografia y estado de las poblaciones y se ha podido comprobar que por suerte, el riesgo de amenaza y extinción se ha minimizado considerablmente al comprobar que el número de individuos y el estado de conservación de las poblaciones es mejor del que se pensaba anteriormente. Por ello, el grado de amenaza ha disminuido y con ello las figuras de protección. Sin embargo, el hecho de que determinadas poblaciones puedan verste afectadas por el aumento de las temperaturas implica la importancia de mantener y proteger las áreas más vulnerables y así intentar minimizar la pérdida de diversidad genética de estas especies de cara al futuro.

Referencias

Barrón E. 2003. Evolución de las floras terciarias en la Península Ibérica. *Monografías del Jardín Botánico de Córdoba* 11: 63–74.

Barrón E, Peyrot D. 2006. La vegetación forestal en el Terciario. In: Carrión J, Fernández S, Fuentes N, eds. *Paleoambientes y cambio climático*. Murcia, Spain: Fundación Séneca/Agencia de Ciencia y Tecnología de la Región de Murcia, 56–77.

Beatty GE, Provan J. 2013. Post-glacial dispersal, rather than in situ glacial survival, best explains the disjunct distribution of the Lusitanian plant species *Daboecia cantabrica* (Ericaceae). *Journal of Biogeography* 40: 335–344.

Beatty GE, Provan J. 2014. Phylogeographical analysis of two cold-tolerant plants with disjunct Lusitanian distributions does not support in situ survival during the last glaciation. *Journal of Biogeography* 41: 2185–2193.

Ben-Menni Schuler S, García-López MDC, Lopez-Flores I, Nieto-Lugilde M, Suárez-Santiago VN. 2017. Genetic diversity and population history of the Killarney fern, *Vandenboschia speciosa* (Hymenophyllaceae), at its southern distribution limit in continental Europe. *Botanical Journal of the Linnean Society* 183: 94–105.

Durães CJ. 2005. Genetic variation and evolutionary history of the Macaronesian fern *Asplenium hemionitis* L. PhD thesis, Edinburgh University, Edinburgh, UK.

Hunt HV, Ansell SW, Russell SJ, Schneider H, Vogel JC. 2009. Genetic diversity and phylogeography in two diploid ferns, *Asplenium fontanum* subsp. *fontanum* and *A. petrarchae* subsp. *bivalens*, in the western Mediterranean. *Molecular Ecology* 18: 4940–4954

Jermy AC. 1984. Origin and distribution of Pteridophytes in the Mediterranean area. *Webbia* 38: 397–416.

Kondraskov P, Schütz N, Schüßler C, de Sequeira MM, Guerra AS, Caujapé-Castells J, Jaén-Molina R, Marrero-Rodríguez A, Koch MA, Linder P *et al*. 2015. Biogeography of Mediterranean Hotspot Biodiversity: Re-Evaluating the 'Tertiary Relict' Hypothesis of Macaronesian Laurel Forests. *PLoS One* 10: e0132091.

Mai DH. 1989. Development and regional differentiation of the European vegetation during the Tertiary. *Plant Systematics and Evolution* 162: 79–91.

Mai DH. 1991. Palaeofloristics changes in Europe and the confirmation of the Arctotertiary–Palaeotropical geofloral concept. *Review of Palaoeobotany and Palynology* 68: 29–36.

Pichi-Sermolli REG. 1979. A survey of the pteridological flora of the Mediterranean region. *Webbia* 34: 175–242.

Pichi Sermolli REG. 1991. On the taxonomy and nomenclature of some species from tropical Africa of the genus *Triplophyllum* Holttum (Dryopteridaceae). *Webbia* 45: 117–135.

Ramil-Rego P, Gómez-Orellana L. 1996. Dinámica climática y biogeográfica del área litoral-sublitoral de Galicia. In: Ramil Rego P, Fernández Rodríguez C, Rodríguez Guitián M, eds. *Biogeografía Pleistocena-Holocena de la Península Ibérica*. Santiago de Compostela, Spain: Xunta de Galicia, 43–71.

Rumsey FJ, Sheffield E. 1996. Inter-generational ecological niche separationand the “independent gametophyte” phenomenom. In: Camus JM, Gibby M, Johns RJ, eds. *Pteridology in perspective*. London, UK: Royal Botanic Gardens, Kew, 563–570.

Rumsey FJ, Vogel JC, Russell SJ, Barrett JA, Gibby M. 1999. Population structure and conservation biology of the endangered fern *Trichomanes speciosum* Willd. (Hymenophyllaceae) at its northern distributional limit. *Biological Journal of the Linnean Society* 66: 333–344.

Van Vuuren DP, Edmonds J, Kainuma M, Riahi K, Thomson A, Hibbard K, Hurtt GC, Kram T, Krey V, Lamarque J-F et al. 2011. The representative concentration pathways: an overview. *Climatic Change* 109: 5.

Vogel JC, Rumsey FJ, Schneller JJ, Barrett JA, Gibby M. 1999a. Where are the glacial refugia in Europe? Evidence from pteridophytes. *Biological Journal of the Linnean Society* 66: 23–37.

Conclusions

1- Our results show a Tertiary origin for *Vandenboschia speciosa, Culcita macrocarpa* and *Diplazium caudatum,* being probable that they have belonged to the Paleotropical Geoflora that lived in Europe from the Late Cretaceous to the Late Miocene along the Thethys Sea. The biogeographical pattern they show seem to have been modeled by changes occured since then, as the geological and climate events of the late Miocene/early Pliocene. This changes shifted Tertiary fern distribution pattern in Europe.

2- The four studied species show medium-high genetic diversity levels, and the highest levels are found in the Cantabrian cornice (for *V. speciosa* and *C. macrocarpa*) and in the Macaronesian archipelagos (especially Azores and Madeira).

3- Population history in Andalusian populations of *V. speciosa* is characterized by a migration-drift equilibrium with historical dispersals as the main factors influencing population structure.

4- *Vandenboschia speciosa* and *C. macrocarpa* share a global phylogeographic pattern characterized by the differentiation of populations in two main groups, North and South. Among the Macaronesian archipelagos, the populations of the Canary Islands and Madeira were related to the populations of the southern group, while the Azores populations were related to those of the north in *V. speciosa* and those of the south in *C. macrocarpa*. *D. caudatum* is characterized by the relation of populations from Azores and the Canary Island, being Madeira more differentiated and Andalusia genetically completely different.

5- The Cantabrian Cornice must have been a main refuge for *V. speciosa* and *C. macrocarpa* and the archipielagos of Macaronesia for the four species during the Tertiary and the Pleistocene glaciations due to the high diversity and private haplotypes they present.

6- *V. speciosa* and *C. macrocarpa* seem to have survived the Pliocene and Pleistocene glaciations in continental Europe, however, *D. caudatum* seems to have disappeared, taking refuge in Macaronesia and subsequently recolonizing the continent.

7- In *V. speciosa*, the Cantabrian Cornice could also have been a recent contact zone between the two lineages due to expansions of the area from the north towards the south and from the Cantabrian towards the north following the glacial-interglacial cycles. In *V. speciosa* current central European populations appear to have derived from post-glacial multiple long-distance dispersals. Macaronesia for both *C. macrocarpa* and *D. caudatum* and the Cantabrian Cornice for *C. macrocarpa* seem to have served as a source for subsequent dispersions and / or recolonizations.

8- The model that best explains the current distribution of the gene lineages observed in *D. caudatum* is long-distance dispersion with restriction of gene flow and genetic drift. *C. macrocarpa* shows a population structure that responds to a dynamic of colonization in which recent populations have originated from one or a few genets, from relict populations, which have been established by intragametophytic selfing and vegetatively expanded. *V. speciosa* biogeographic pattern is the result of long persistence in independent refugia, independent evolution of gene pools and posterior North-South dispersals tracking the Quaternary glacial cycles.

9- In *Pteris incompleta* we characterized eight microsatellite loci enabling biogeographic, phylogeographic and genetic diversity studies of this relict fern.

Conclusiones

1- Nuestros resultados muestran un origen terciario de *Vandenboschia speciosa*, *Culcita macrocarpa* y *Diplazium caudatum*, siendo probable que hayan pertenecido a la Geoflora Paleotropical que vivió en Europa desde el Cretácico tardío hasta el Mioceno tardío a lo largo del mar de Thethys. El patrón biogeográfico que muestran parece haber sido modelado por los cambios ocurridos desde entonces, como los eventos geológicos y climáticos del Mioceno tardío / Plioceno temprano. Estos cambios desplazaron el patrón de distribución de helechos terciarios en Europa.

2- Las cuatro especies estudiadas muestran niveles de diversidad genética medio-altos, y los niveles más altos se encuentran en la Cornisa Cantábrica (para *V. speciosa* y *C. macrocarpa*) y en los archipiélagos macaronesios (especialmente Azores y Madeira).

3- La historia poblacional de las poblaciones andaluzas de *V.* speciosa se caracteriza por un equilibrio migración-deriva génica con dispersiones históricas como los principales factores que influyen en la estructura de las poblaciones.

4- *V. speciosa* y *C. macrocarpa* comparten un patrón filogeográfico global caracterizado por la diferenciación de las poblaciones en dos grupos principales, Norte y Sur. Entre los archipiélagos macaronesios, las poblaciones de las Islas Canarias y Madeira estan relacionadas con las poblaciones del grupo del sur, mientras que las poblaciones de las Azores estan relacionadas con las del norte en *V. speciosa* y las del sur en *C. macrocarpa*. *D. caudatum* se caracteriza por la relación entre las poblaciones de Azores y Canarias, estando Madeira más diferenciada y siendo Andalucía genéticamente diferente al resto.

5- La Cornisa Cantábrica debe haberse comportado como refugio principal para *V. speciosa* y *C. macrocarpa* y los archipiélagos macaronésicos para las cuatro especies durante el Terciario y las glaciaciones del Pleistoceno debido a la alta diversidad y haplotipos privados que presentan.

6- *V. speciosa* y *C. macrocarpa* parecen haber sobrevivido al Plioceno y las glaciaciones del Pleistoceno en Europa continental, sin embargo, *D. caudatum* parece haber desaparecido, refugiándose en la Macaronesia y posteriormente recolonizando el continente.

7- En *V. speciosa*, la Cornisa Cantábrica también podría haber sido una zona de contacto reciente entre los dos linajes debido a las expansiones del área desde el norte hacia el sur y desde el Cantábrico hacia el norte siguiendo los ciclos glaciales-interglaciales. Además, las poblaciones actuales de Europa central de *V. speciosa* parecen haberse derivado de múltiples dispersiones postglaciales a larga distancia. La Macaronesia para *C. macrocarpa* y *D. caudatum* y la Cornisa Cantábrica para *C. macrocarpa* parecen haber servido como fuente de dispersiones y/o recolonizaciones posteriores.

8- El modelo que mejor explica la distribución actual de los linajes génicos observados en *D. caudatum* es la dispersión a larga distancia con restricción del flujo génico y la deriva genética. *C. macrocarpa* muestra una estructura poblacional que responde a una dinámica de colonización en la que las poblaciones recientes se han originado a partir de uno o unos pocos genets, a partir de poblaciones relictas, que se han establecido mediante la autofecundación intragametofítica y expandido vegetativamente. El patrón biogeográfico de *V. speciosa* es el resultado de la larga persistencia en refugios independientes, la evolución independiente de los pools génicos y dispersiones posteriores Norte-Sur siguiendo los ciclos glaciales cuaternarios.

9- En *Pteris incompleta* caracterizamos ocho loci de microsatélites que permiten estudios biogeográficos, filogeográficos y de diversidad genética de este helecho relicto.