



FACULTADE DE CIENCIAS

GRAO EN BIOQUÍMICA

D. Carlos Daniel Mirás Viña

Aplicación de marcadores de ADN para análisis de diversidad genética de la especie *Daphne gnidium* en poblaciones costeras atlánticas de Galicia.

Aplicación de marcadores de ADN para a análise da diversidade xenética da especie *Daphne gnidium* en poboacións costeiras atlánticas de Galicia.

Application of DNA markers for genetic diversity analysis of the species *Daphne gnidium* in Atlantic coastal populations of Galicia.

Traballo Fin de Grao

Xuño 2024

ÍNDICE

RESUMEN	4
RESUMO	5
ABSTRACT	6
1. INTRODUCCIÓN	7
1.1 <i>Daphne gnidium</i>	7
1.2 Marcadores moleculares	11
1.3 Estudios genéticos previos	14
2. OBJETIVOS	17
3. MATERIAL Y MÉTODOS	17
3.1 Material biológico	17
3.2 Extracción de ADN	18
3.3 Selección y validación de un panel de microsatélites	19
3.4 Caracterización poblacional de marcadores genéticos	24
3.4.1 Análisis de diversidad genética	25
3.4.2 Análisis de diferenciación y estructura genética	26
4. RESULTADOS Y DISCUSIÓN	28
4.1 Extracción de ADN	28
4.2 Amplificación y genotipado de marcadores microsatélite ortólogos	28
4.3 Análisis de resultados de marcadores microsatélites	32
4.3.1 Frecuencias alélicas	32
4.3.2 Análisis de diversidad genética	33
4.3.3 Análisis de diferenciación y estructura genética	38
4.4 Análisis poblacional con marcadores SNPs	40
4.4.1 Análisis de diversidad genética	40
4.4.2 Análisis de diferenciación y estructura genética	43
5. CONCLUSIONES	47
6. BIBLIOGRAFÍA	48

RESUMEN

El presente trabajo fin de grado (TFG) se centra en la aplicación de marcadores de ADN para evaluar la diversidad y diferenciación genética de la especie *Daphne gnidium* en áreas costeras insulares protegidas del Parque Nacional Illas Atlánticas y áreas continentales atlánticas de Galicia (Rías Baixas), donde no existen datos genéticos previos.

Daphne gnidium (familia Thymelaeaceae), es un arbusto perenne con interés medicinal y de conservación, principalmente distribuida en el Mediterráneo. Ante la ausencia de marcadores específicos de ADN para esta especie, el análisis poblacional se abordó mediante la puesta a punto de un panel de marcadores microsatélite, previamente descritos en la especie congénérica *D. rodriguezii*, y a través del estudio de polimorfismos de un solo nucleótido (SNPs) obtenidos por tecnologías de genotipado.

En las áreas de estudio se recolectaron muestras foliares de aproximadamente 20 plantas por localidad. Se realizó la extracción de ADN y la amplificación cruzada de 9 loci microsatélite aislados en *D. rodriguezii*, optimizando las condiciones de amplificación *multiplex*, aunque con baja tasa final de loci informativos. Además, se realizó la caracterización de un panel de marcadores de tipo SNP generados mediante análisis genómicos en *D. gnidium*.

Los marcadores seleccionados se analizaron para evaluar los niveles de diversidad y diferenciación genética entre las poblaciones continentales e insulares. Los resultados mostraron una distribución heterogénea de la diversidad, con mayores valores en algunas localidades respecto a otras menos variables, sugiriendo efectos de aislamiento y deriva genética. El análisis de SNPs indicó una mayor diferenciación interpoblacional y estructuración diferenciando dos grupos de localidades insulares en relación al área continental.

Estos resultados permitieron obtener por primera vez datos sobre la diversidad y estructura genética de *D. gnidium* en las localidades estudiadas, proporcionando información útil para

su conservación y gestión, en el marco de un programa de restauración de hábitats insulares atlánticos.

Palabras clave: *Daphne gnidium*, marcadores genéticos, microsatélites, SNP, diversidad genética, diferenciación genética.

RESUMO

O presente traballo de fin de grado (TFG) céntrase na aplicación de marcadores de ADN para avaliar a diversidade e diferenciación xenética da especie *Daphne gnidium* en áreas costeiras illeiras protexidas do Parque Nacional Illas Atlánticas e áreas continentais atlánticas de Galicia (Rías Baixas), onde non existen datos xenéticos previos.

Daphne gnidium (familia Thymelaeaceae) é un arbusto perenne con interese medicinal e de conservación, principalmente distribuído no Mediterráneo. Ante a ausencia de marcadores específicos de ADN para esta especie, a análise poboacional abordouse mediante a posta a punto dun panel de marcadores microsatélite, previamente descritos na especie conxénere *D. rodriguezii*, e a través do estudo de polimorfismos dun só nucleótido (SNPs) obtidos por tecnoloxías de xenotipado.

Nas áreas de estudo recolléronse mostras foliares de aproximadamente 20 plantas por localidade. Realizouse a extracción de ADN e a amplificación cruzada de 9 loci microsatélite illados en *D. rodriguezii*, optimizando as condicións de amplificación multiplex, aínda que cunha baixa taxa final de loci informativos. Ademais, realizouse a caracterización dun panel de marcadores de tipo SNP xerados mediante análises xenómicas en *D. gnidium*.

Os marcadores seleccionados analizáronse para avaliar os niveis de diversidade e diferenciación xenética entre as poboacións continentais e illeiras. Os resultados mostraron unha distribución heteroxénea da diversidade, con maiores valores nalgúns lugares respecto a outros menos variables, suxirindo efectos de illamento e deriva xenética.

A análise de SNPs indicou unha maior diferenciación interpoblacional e estruturación que diferenciaba dous grupos de lugares illeiros en relación á área continental.

Estes resultados permitiron obter por primeira vez datos sobre a diversidade e estrutura xenética de *D. gnidium* nos lugares estudados, proporcionando información útil para a súa conservación e xestión, no marco dun programa de restauración de hábitats atlánticos illeiros.

Palabras clave: *Daphne gnidium*, marcadores xenéticos, microsátélites, SNP, diversidade xenética, diferenciación xenética.

ABSTRACT

The present bachelor's thesis (TFG) focuses on the application of DNA markers to evaluate the genetic diversity and differentiation of the species *Daphne gnidium* in the protected coastal insular areas of the Atlantic Islands National Park and the Atlantic continental areas of Galicia (Rías Baixas), where there are no previous genetic data.

Daphne gnidium (Thymelaeaceae family), is a perennial shrub with medicinal and conservation interest, mainly distributed in the Mediterranean region. Due to the absence of specific DNA markers for this species, the population analysis was approached by developing a panel of microsatellite markers, previously described in the congeneric species *D. rodriguezii*, and through the study of single nucleotide polymorphisms (SNPs) obtained using genotyping technologies.

Leaf samples were collected from approximately 20 plants per locality in the study areas. DNA extraction and cross-amplification of 9 microsatellite loci isolated in *D. rodriguezii* were performed, optimizing the multiplex amplification conditions, although with a low final rate of informative loci. Additionally, a panel of SNP-type markers generated through genomic analyses in *D. gnidium* was characterized.

The selected markers were analyzed to evaluate the levels of genetic diversity and differentiation between the continental and insular populations. The results showed a heterogeneous distribution of diversity, with higher values in some localities compared to others with less variability, suggesting effects of isolation and genetic drift. The SNP analysis indicated a greater interpopulation differentiation and structuring, differentiating two groups of insular localities in relation to the continental area.

These results allowed obtaining, for the first time, data on the genetic diversity and structure of *D. gnidium* in the studied localities, providing useful information for its conservation and management, within the framework of a restoration program for Atlantic insular habitats.

Keywords: *Daphne gnidium*, genetic markers, microsatellites, SNP, genetic diversity, genetic differentiation.

1. INTRODUCCIÓN

1.1 *Daphne gnidium*

La especie *Daphne gnidium*, denominada torvisco, es un arbusto perteneciente a la familia de las timeláceas, la cual comprende unos 50 géneros y 500 especies. Su encuadre taxonómico es el siguiente:

- Reino: Plantae
 - Filo: Estreptofita
 - Clase: Magnoliopsida
 - Orden: Malvales
 - Familia: Thymelaeaceae
 - Género: *Daphne*
 - Especie: *Daphne gnidium*

Existen numerosas discrepancias respecto al número de especies que comprende el género *Daphne*, consecuencia de una falta de un estudio filogenético integral del género, del elevado número de especies y de la escasez de caracteres morfológicos distintivos para delimitar las distintas especies (Banerjee *et al.*, 2022). Dicha diversidad es destacable en

Asia, especialmente en China. Por otra parte, Europa alberga unas diecisiete especies, diez de las cuales se distribuyen geográficamente en el Mediterráneo incluyendo la especie objeto de estudio, *D. gnidium* (Ramos-Capón *et al.*, 2024).

En *D. gnidium* se reconocen tres subespecies: *D. gnidium subsp. gnidio*, *D. gnidium subsp. mauritanica* y *D. gnidium subsp. maritima*, cada una de las cuales con características diferentes y distribuciones geográficas específicas. No obstante, la circunscripción infraespecífica del género, al ser polifilética, sigue siendo motivo de debate entre los taxonomistas (Ramos-Capón *et al.*, 2024).

D. gnidium es una planta ampliamente distribuida a distinta altitud en áreas suratlánticas mediterráneas, en el sur de Europa, en áreas templadas y subtropicales de Asia, y el Norte de África (Fig. 1). Se extiende desde la Península Ibérica hasta Grecia, incluyendo las Islas Canarias y, en mayor medida, las Islas Baleares. En estas últimas, la especie puede encontrarse principalmente en hábitats abiertos de las zonas oeste y central del archipiélago. Crece en regiones áridas con escasas precipitaciones y a menudo se encuentra en las zonas de matorral y sotobosque y en las zonas costeras. Presenta gran capacidad de revegetación de áreas dañadas como incendios o erosión (Khouchlaa *et al.*, 2021).

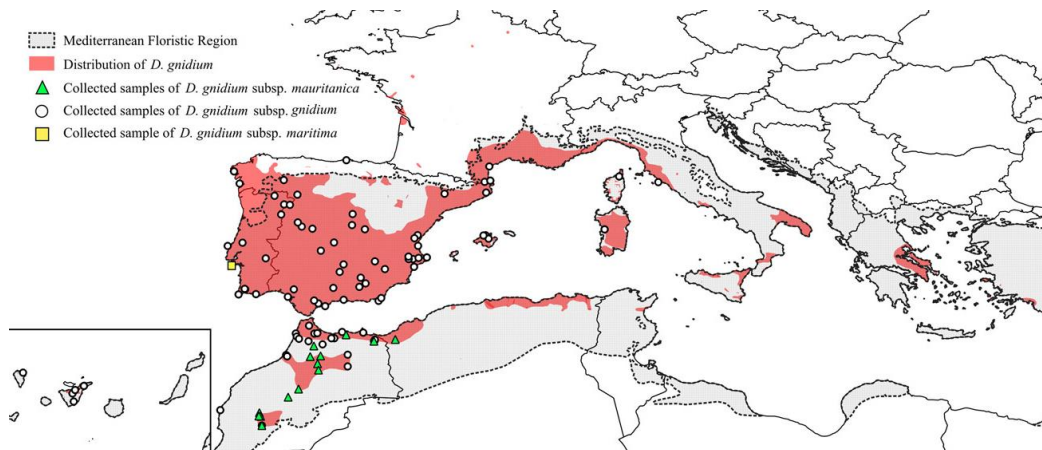


Figura 1. Rango de distribución de *D. gnidium* (magenta). Fuente: Ramos-Capón *et al.* (2024).

El torvisco es una especie diploide ($2n=18$), de interés medicinal y de conservación. Es conocida popularmente como ‘matagallinas’ porque su corteza produce una especie de resina que, al contacto con la piel o las mucosas, provoca una fuerte irritación y sus frutos, una especie de baya, son altamente tóxicos.

D. gnidium es un arbusto que puede alcanzar los dos metros de altura, de tipo esclerófilo, es decir, se caracteriza por la presencia de hojas pequeñas, perennes, gruesas, duras y coriáceas, permitiendo adaptarse a climas secos y cálidos (Fig. 2). La dureza de sus hojas se debe a una cutícula gruesa y a una epidermis con células escleróticas. Sus flores hermafroditas son tubulares y de color crema localizadas en la parte apical de las nuevas ramas con un olor dulce y son polinizados por insectos. Florece de julio a noviembre y sus flores tienen una vida corta de unos 3-4 días. Los frutos son bayas rojo-anaranjadas, carnosas, con una sola semilla. Maduran hasta diciembre y cada arbusto produce un promedio de 250 frutos (Khouchlaa *et al.*, 2021).

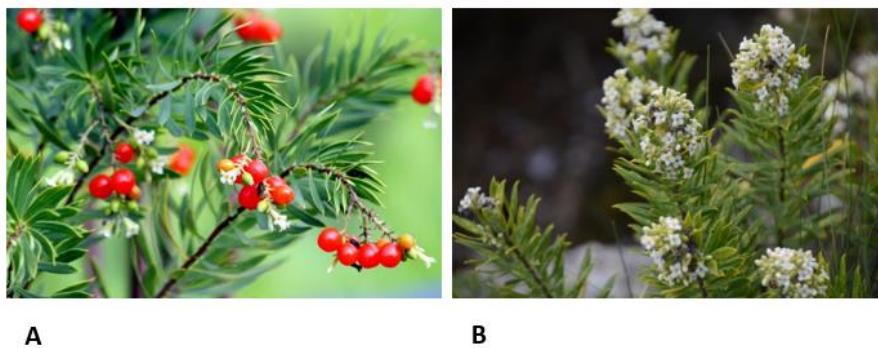


Figura. 2. A: Bayas de *Daphne gnidium* L. Fuente: © Pablo Vargas | RJB-CSIC . **B:** Planta florada de *Daphne Gidium* L. Fuente: LIFE INSULAR, IBADER.

Los estudios etnobotánicos han resaltado la relevancia de *D. gnidium* en la medicina tradicional, debido a los diversos usos terapéuticos de sus diferentes partes, que varían según la ubicación geográfica de la planta. Estos usos medicinales tradicionales resaltan la importancia de conservar y estudiar esta especie vegetal con interés farmacológico.

Las hojas son la parte más utilizada, con propiedades farmacológicas, médicas y cosméticas específicas según la región de origen. En Marruecos se emplean para el cuidado

del cabello y, junto con la corteza, se utilizan ante problemas dermatológicos y como hipoglucemiantes para tratar la diabetes. En Argelia, se usa ante crisis nerviosas, contra dolores reumáticos y musculares, y como antiinflamatorio. En Túnez las hojas se prescriben para tratar el cáncer de piel mientras que en Italia se utilizan como antiparasitario y remedio gastrointestinal. Además, en líneas generales, las partes aéreas poseen propiedades antiinflamatorias útiles para la odontalgia. Por otra parte, en España, las raíces y ramas se usan como insecticida, las flores para enfermedades cutáneas, y los frutos contra infecciones gastrointestinales (Khouchlaa *et al.*, 2021).

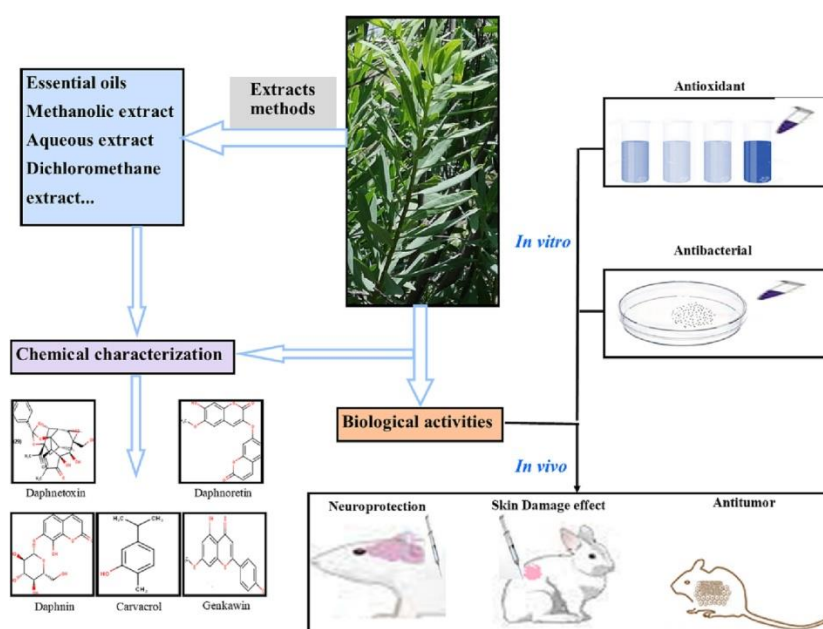


Figura 3. Representación de algunas actividades biológicas probadas por diferentes extractos de *D. gnidium*. Fuente: El Mouzazi *et al.* (2023).

Investigaciones farmacológicas han demostrado que la presencia de ciertos compuestos fitoquímicos en aceites esenciales y extractos de la especie son responsables de los efectos biológicos de la especie, sugiriendo su uso como fuente de moléculas interesantes para aplicaciones futuras en medicina y cosmética. Algunos de estos efectos son: antifúngico, antibacteriano, antigenotóxico, antioxidante, antitumoral, antiviral, hemolítico, antiinflamatorio, inmunomodulador, neuroprotector, contra el daño epidérmico, alelopático, insecticida y efectos inhibitorios de corrosión (Fig. 3). Los compuestos

fitoquímicos son sustancias químicas naturales responsables del color, sabor, aroma y textura que presentan las plantas. Son variables en función del origen de la planta, lo que puede explicar las diferentes aplicaciones farmacológicas entre distintos ejemplares (El Mouzazi *et al.*, 2023). Algunos de estos compuestos presentes en *D. gnidium* son terpenoides, flavonoides, cumarinas, ácidos grasos y alcanos, destacando como grupo principal los flavonoides (Fig.4).

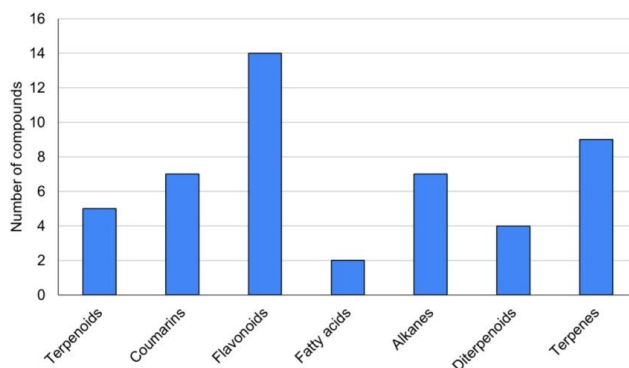


Figura 4. Principales compuestos bioactivos encontrados en *D. gnidium*. Fuente: El Mouzazi *et al.* (2023).

1.2 Marcadores moleculares

El desarrollo de marcadores moleculares ha sido fundamental para los estudios de genética poblacional, ya que permiten investigar la variabilidad genética y la endogamia, entre otras aplicaciones. Los marcadores moleculares de ADN son los más utilizados, correspondiéndose a segmentos genómicos ubicados en cromosomas (loci) y cuya herencia se puede rastrear, representando una de las herramientas más eficaces para el estudio de la variación genética y la diversidad biológica. Estos marcadores pueden ser de herencia dominante o codominante (permite diferenciar entre individuos homocigotos y heterocigotos) y neutrales (sus variantes no suponen un cambio fenotípico o metabólico) o no neutrales (producen un cambio y, por lo tanto, sometido a selección). Además, factores como su abundancia, reproducibilidad, polimorfismo y bajo coste han contribuido a la creciente difusión y utilización de estos marcadores en diversas investigaciones (Pierce *et al.*, 2021).

Los marcadores moleculares de ADN más utilizados en los últimos años son los microsatélites, también conocidos como SSRs (Repetición de Secuencias Simples; Jacob *et al.*, 1991) o STRs (Repeticiones Cortas en Tándem; Craig, 1993), junto con los polimorfismos de un solo nucleótido (*single nucleotide polymorphism*, SNP). Estos marcadores se utilizan con frecuencia en genética de poblaciones para estudiar la estructura y la dinámica de estas, ya sea para análisis de diversidad genética entre individuos de una misma población, el flujo genético entre poblaciones de la misma especie o para la identificación individual y construcción de la historia evolutiva de una especie.

Los marcadores microsatélites son regiones del ADN que contienen repeticiones en tándem de secuencias cortas (de 1 a 6 nucleótidos). Se caracterizan por su herencia codominante y su elevado polimorfismo debido a su alta tasa de mutación bastante variable (Zimmerman *et al.*, 2020), entre 10^{-2} y 10^{-5} por locus y por generación (Ellegren, 2000). Además, pueden ser fácilmente analizados mediante la Reacción en Cadena de la Polimerasa (PCR) incluso partiendo de ADN degradado, a un coste muy bajo, son muy reproducibles, ampliamente transferibles entre especies relacionadas e incluso entre géneros al presentar regiones adyacentes altamente conservadas y haber acceso a las secuencias de los cebadores utilizados para su amplificación mediante PCR en bases de datos públicas (Pierce *et al.*, 2021).

La principal limitación de los marcadores microsatélites radica en la necesidad de conocer previamente la secuencia de nucleótidos que los flanquea, a fin de poder diseñar los cebadores para su amplificación. En especies sin un conocimiento genómico previo es necesario aislarlos *de novo* mediante librerías enriquecidas en estos motivos (Zane *et al.*, 2002) mediante tecnologías de nueva secuenciación o utilizar marcadores aislados en especies relacionadas para testar su amplificación cruzada (García-Verdugo *et al.*, 2019). Asimismo, el uso de distintos equipos y la subjetividad en el establecimiento de los tamaños alélicos entre diferentes laboratorios puede afectar a la estandarización del genotipado, lo que constituye otro desafío en el uso de estos marcadores (Zimmerman *et al.*, 2020).

Los marcadores SNPs son variaciones de un solo nucleótido. Dado que el ADN contiene cuatro nucleótidos (adenina, guanina, citosina y timina), un SNP como máximo tendrá cuatro alelos. No obstante, la mayoría de los SNPs son bialélicos y los más frecuentes son los que corresponden al cambio de una base púrica por otra o de una pirimidínica por otra, A/G o C/T, respectivamente. La reducida diversidad alélica por locus de los SNPs determina que sean menos informativos a nivel individual frente a los microsatélites, en general altamente polimórficos (Zimmerman *et al.*, 2020), por lo que se necesitan mayores conjuntos de SNPs para obtener unos resultados con alto potencial informativo (García-Verdugo *et al.*, 2019). Los SNPs son las variaciones genéticas más comunes ampliamente distribuidos a través del genoma.

Debido a la abundancia y estabilidad de estos marcadores, las estimas de los parámetros poblacionales son más precisas que con microsatélites, lo que se traduce en un mejor ajuste del nivel de diversidad poblacional (Brookes, 1999). Además, el desarrollo de técnicas genómicas ha facilitado la detección y genotipado de grandes conjuntos de SNPs en especies no modelo lo que ha extendido su uso en los últimos años (Pérez-González *et al.*, 2023). Las técnicas de genotipado por secuenciación que aplican la secuenciación masiva a librerías de representación reducida, se basan en la secuenciación de un pequeño porcentaje de ADN asociada al sitio de restricción (*Restriction-site Associated DNA sequencing*; RAD-seq) mediante el uso de endonucleasas para fragmentar el genoma. La amplificación por PCR y la secuenciación de estos fragmentos generados permite la identificación y genotipado de miles de loci con SNPs útiles para el análisis genético poblacional sin necesidad de conocimiento genómico previo (Chambers *et al.*, 2023).

Por lo tanto, aunque los resultados obtenidos con distintos tipos de marcadores pueden resultar similares, los SNPs se presentan como una alternativa para obtener una evaluación más precisa de la diversidad y diferenciación genética a escala genómica de la especie de interés (Zimmerman *et al.*, 2020).

1.3 Estudios genéticos previos

Estudios filogenéticos previos describen la incertidumbre taxonómica a nivel infraespecífico donde el número de especies del género *Daphne* oscila entre las 42 especies reconocidas en los años 70 y las 97 recientemente aceptadas (POWO, 2024). Esta falta de consenso se debe principalmente a la ausencia de un estudio filogenético integral del género y a la escasez de caracteres morfológicos claros para delimitar las especies (Ramos-Capón *et al.*, 2024).

A pesar de ello, *D. gnidium*, al ser una especie que muestra rasgos del síndrome esclerófilo (hojas coriáceas, flores discretas, frutos carnosos), se considera un candidato apropiado para evaluar la hipótesis de que los linajes con este síndrome experimentaron bajas tasas de especiación tras el establecimiento del clima mediterráneo (Rubio de Casas, 2006; Ramos-Capón *et al.*, 2024).

En Ramos-Capón *et al.* (2024) se realizó un análisis filogenético a nivel de género *Daphne* y de la familia Thymelaeaceae utilizando secuencias de ADN nuclear (ITS: *Internal Transcribed Spacer*) y cloroplastídico (trnV: *transfer RNA of Valine*, rps16: *small ribosomal proteins*). Los resultados revelaron que la circunscripción actual del género es polifilética, con algunas especies anidadas dentro de otros géneros. Además, el análisis filogeográfico demostró una baja diferenciación genética de *D. gnidium gnidium* y *D. gnidium marítima* a pesar de su amplio rango de distribución, apoyando la hipótesis del síndrome esclerófilo sin signos de especiación tras la implantación del clima mediterráneo.

El análisis de los recursos genéticos desempeña un papel fundamental en la evaluación y conservación de la diversidad biológica. El conocimiento de los niveles de diversidad intra e interpoblaciones y su distribución geográfica, es fundamental para diseñar estrategias adecuadas de manejo y conservación (Bouza *et al.*, 2023). Estos programas se centran en el análisis de diversidad genética y caracterización de los recursos genéticos (García-Verdugo *et al.*, 2015). Esto es importante en los límites de la distribución de las especies,

donde las poblaciones suelen presentar una menor diversidad genética y mayor divergencia con respecto a poblaciones centrales (Bouza *et al.*, 2023).

Asimismo, la disminución del tamaño efectivo de las poblaciones y la pérdida de conectividad pueden derivar en una disminución de la diversidad genética, haciéndolas más susceptibles a los efectos de la endogamia y deriva genética (Bouza *et al.*, 2023). Por lo tanto, es necesario considerar estos factores al implementar medidas de conservación, como las translocaciones de individuos, para evitar riesgos como la saturación genética o la pérdida de genotipos únicos.

Se ha descrito de forma generalizada que las poblaciones insulares presentaban una menor diversidad genética en comparación con las poblaciones continentales. Esta observación se ha relacionado con el aislamiento geográfico de las islas que favorece la pérdida de variabilidad genética. Sin embargo, estudios recientes han revisado esta hipótesis, revelando que existen múltiples factores que influyen en los niveles de diversidad genética de las poblaciones insulares (García-Verdugo *et al.*, 2015). Los análisis realizados por García-Verdugo *et al.* (2015) sobre *Periploca laevigata* en el archipiélago canario, no detectaron la existencia de un patrón generalizado que mostrara diferencias significativas en los niveles de diversidad genética entre poblaciones insulares y continentales. Esto sugiere que, si bien no se puede establecer una regla general, algunas poblaciones insulares pueden presentar una variación genética menor que las continentales.

Las comparaciones directas de estimaciones de diversidad genética entre poblaciones continentales e insulares se han convertido en el enfoque más utilizado para evaluar los niveles de diversidad genética en las islas. Estas comparaciones permiten identificar los factores que modulan la diversidad, como la capacidad de dispersión de las especies o la distancia al continente (García-Verdugo *et al.*, 2015).

Los efectos de los factores ambientales suelen ser más determinantes en las poblaciones aisladas, como es el caso de las especies insulares. Debido a que estas poblaciones presentan una menor diversidad genética y un mayor grado de diferenciación, se vuelven

más susceptibles a los cambios ambientales (Bouza *et al.*, 2023). No obstante, esta situación también puede dar lugar al desarrollo de adaptaciones únicas, lo que resalta la importancia de estas poblaciones para la conservación.

En el caso de la especie *D. gnidium*, se ha estudiado su biología reproductiva con el objetivo de identificar los factores responsables de la baja producción de frutos observada (Roccoliello *et al.*, 2009). Los resultados sugieren que los polinizadores no desempeñan un papel crucial en el sistema de reproducción de esta especie, y que el desarrollo del gametofito femenino puede detenerse en diferentes etapas, posiblemente como un mecanismo para evitar la autopolinización. Estos hallazgos apuntan a un proceso evolutivo hacia una andromonoecia o androdioicismo funcional en *D. gnidium*, aunque se requieren más investigaciones para confirmar esta hipótesis.

Adicionalmente, en un estudio sobre la especie congénérica, *Daphne rodriguezii*, se han caracterizado un conjunto de marcadores moleculares microsatélite (García-Verdugo *et al.*, 2019). Estos marcadores han demostrado ser una herramienta poderosa y rentable para detectar niveles sustanciales de variación genética, incluso en poblaciones pequeñas. Además, se ha evaluado la amplificación cruzada de estos loci en especies estrechamente relacionadas, como *D. gnidium*, indicando que estos marcadores pueden ser fácilmente aplicados en estudios genéticos poblacionales de otras especies del género *Daphne*.

El análisis de los recursos genéticos desempeña un papel fundamental en la evaluación y conservación de la diversidad biológica. Aunque la hipótesis de una menor diversidad genética en las poblaciones insulares no se ha confirmado de manera generalizada, existen casos en los que estas poblaciones presentan niveles más bajos de diversidad. La comprensión de los factores que modulan la diversidad genética, tanto en las islas como en los continentes, es crucial para diseñar estrategias de conservación efectivas y garantizar la supervivencia a largo plazo de las especies.

2. OBJETIVOS

El presente trabajo fin de grado se desarrolla en el marco de un proyecto de restauración de hábitats insulares en el océano Atlántico. En este contexto, este estudio se centra en el análisis genético de poblaciones de *Daphne gnidium* ubicadas en áreas continentales de las Rías Baixas gallegas y en áreas costeras insulares protegidas, como el Parque Nacional Illas Atlánticas, donde no existen datos genéticos previos sobre esta especie.

En este contexto los objetivos que se han planteado en el presente trabajo han sido:

1. Puesta a punto de un panel de marcadores microsatélites mediante amplificación cruzada.
2. Caracterización de un amplio conjunto de marcadores SNP obtenidos mediante análisis genómico.
3. Realizar un análisis de diversidad, diferenciación genética y estructura poblacional, mediante marcadores microsatélites y SNPs, en muestras poblacionales de áreas costeras insulares y continentales atlánticas de Galicia, para aportar/generar información útil de cara a apoyar acciones de gestión y conservación de los recursos genéticos de *Daphne gnidium*.

3. MATERIAL Y MÉTODOS

3.1 Material biológico

El material biológico empleado para llevar a cabo el presente estudio de *D. gnidium* se ha obtenido en un muestreo representativo aleatorio en 3 localidades insulares protegidas (Parque Nacional de Illas Atlánticas de Galicia) y en 3 áreas continentales atlánticas adyacentes (Rías Baixas) (Tabla 1). El tamaño muestral ha sido de 20 individuos por localidad. El análisis poblacional con marcadores microsatélites se realizó en todas las localidades muestreadas, mientras que para los marcadores SNPs se analizaron las muestras disponibles genotipadas por secuenciación en 4 localidades (Tabla 1). Los muestreos en este estudio se enmarcan en el proyecto LIFE-INSULAR

(<https://www.lifeinsular.eu/>), en colaboración con el Instituto IBADER de la Universidad de Santiago de Compostela (<https://www.ibader.gal/>).

Como material de partida se utilizaron entre 50-100 mg de hojas sanas frescas de cada individuo. Las muestras recogidas en campo se conservaron en sobres individuales identificando cada ejemplar con un código que incluye información de la población de origen (I1, I2, I3, C1, C2, C3) y el número de ejemplar dentro de cada población (Tabla 1). Una vez en el laboratorio, las muestras se desecaron a temperatura ambiente, para evitar su deterioro y contaminación por hongos, permitiendo su almacenamiento a largo plazo.

Tabla 1. Zonas de muestreo y número de ejemplares de la especie *Daphne gnidium* analizados en el presente trabajo.

Localidad	Código genético (Coordenadas)	Número de Muestras*
Cíes	I1 (-8,8992229; 42,2279232)	20 (16)
Ons	I2 (-8,9281983; 42,3866004)	20 (16)
Sálvora	I3 (-9,0038148; 42,4699527)	20 (16)
Barra	C1 (-8,8360239; 42,2628342)	20 (16)
O Grove	C2 (-8,8950985; 42,4733250)	20 (nd)
Corrubedo	C3 (-9,0283086; 42,5491178)	20 (nd)
Total		120 (64)

*Entre paréntesis, tamaño muestral para SNPs, en algún caso no disponible (nd).

3.2 Extracción de ADN

La extracción de ADN de cada muestra individual se realizó siguiendo un protocolo ajustado previamente en otras plantas por el grupo ACUIGEN-USC (Vera *et al.*, 2020; Maroso *et al.*, 2021). Implica la utilización del kit de extracción E.Z.N.A.® Plant DNA DS Kit (Omega) que permite la extracción eficiente de ADN genómico de hasta 30 kb de tamaño a partir de tejido vegetal fresco, congelado o seco, incluso con alto contenido en polisacáridos, polifenoles o bajo contenido de ADN. Se obtiene así, ADN purificado libre de contaminantes.

De cada muestra se pesaron 37 mg de hojas desecadas y se introdujeron en un tubo de 1.5 mL con 700 μ L de tampón CSPL y 20 μ L de Proteinasa K (20 mg/ mL). La muestra se trituró mecánicamente con unas tijeras para facilitar su disolución y se procedió a la extracción del ADN siguiendo las instrucciones del fabricante salvo ciertos cambios, para mejorar el rendimiento y calidad del ADN obtenido, que se indican a continuación.

En primer lugar, en el punto tres del protocolo (*Quick Guide*) se incubó toda la noche a 65°C en lugar de 30 min como indica el protocolo del kit, para que la degradación del tejido fuese lo más eficaz posible. En el paso número seis, la incubación en RNasa A se aumentó de 5 min a 30 min. Por otra parte, en el paso nueve, en lugar de transferir 750 μ L de lisado a través de la columna se transfirieron 600 μ L y se aumentó la centrifugación de 12.000 a 13.000 g. Asimismo, los pasos 8-9 y 11-12 se repitieron tres veces en lugar de dos. Finalmente, en el paso 14 se centrifugó en primer lugar durante 1 min y después durante 2 min a 13.000 g, en lugar de realizar una única centrifugación de 2 min.

El ADN extraído se analizó en un espectrofotómetro NanoDrop ND-1000 para determinar su concentración y calidad en función de los ratios 260/280 (1.8-2,1 pureza óptima) y 260/230 (2,0-2,2 pureza óptima). A continuación, se prepararon alícuotas de trabajo a una concentración de 40 ng/ μ L. Finalmente, el ADN se conservó a 4°C o a -20°C en función de su uso a corto o a largo plazo, respectivamente.

3.3 Selección y validación de un panel de microsatélites

Para el presente estudio se seleccionaron 9 loci microsatélites de los 15 publicados y aplicados en anteriores estudios poblacionales en la especie congénica *D. rodriguezii* (García-Verdugo *et al.*, 2019) ante la ausencia de marcadores microsatélites específicos en *D. gnidium*. La selección de estos 9 marcadores se realizó teniendo en cuenta los resultados descritos por García-Verdugo *et al.* (2019) en la amplificación cruzada en *D. gnidium*.

La puesta a punto inicial de la amplificación de los 9 loci microsatélites se llevó a cabo mediante PCR *singleplex* (amplificación de cada locus microsatélite individualmente), en 3

muestras de ADN de la población de Cíes, utilizando los cebadores publicados para estos loci (Tabla 2) sin marcaje fluorescente.

Tabla 2. Características de los 9 loci microsatélites descritos en *D. rodriguezii* para su amplificación cruzada en *D. gnidium* (modificada de García-Verdugo *et al.*, 2019). pb: pares de base.

Locus	Secuencias de cebadores (5'-3')	Motivo de repetición	Rango alélico <i>D. rodriguezii</i> (pb)	Número de alelos en <i>D. rodriguezii</i>	Rango alélico <i>D. gnidium</i> (pb)	Número de alelos en <i>D. gnidium</i>
Dro-12	F: CATAATGCTGACGTGGACGG	(CTT) 9	237-270	11	255	1
	R: ATGGAGGCGGAAACTTAGG					
Dro-19	F: CGGAGGGATTCAAACCTGGG	(ATT) 9	246-258	5	247	1
	R: TGTTGACTTCAATTTTGTGCGG					
Dro-25	F: TAACGGCATTGCAGGTTTTTC	(TA) _{26}	157-159	2	152-156	3
	R: GGGTGTATAGCCCCTACGTC					
Dro-28	F: TAAAAAGCGCCGGACTCAAG	(TCT) 12	175-190	6	198	1
	R: AGCTGGTTCCTGACGATGAG					
Dro-35	F: AACATCGATTTCTGTGCGGG	(TA) _{16}	204-210	2	192	1
	R: ACGGGGCTTTTGTGCATTC					
Dro-41	F: GAATCCCAACTGCATCGTGG	(GAT) 14 GG(TGA) ₉	240-261	8	184	1
	R: TGGGCTTGTGCATGGTAAACG					
Dro-42	F: AGGGTATTTTCGTGGGCTGG	(TTA) 21	243-318	17	164-168	1
	R: ACAATGTAAAAAGCAAGAAATCCAC					
Dro-46	F: CCCGCTTACTTCAGTGTGC	(CAT) _{12}	180-204	7	173-179	2
	R: ATCGCTAAGATTTCGGGTTGC					
Dro-73	F: GACTGACGATGTCTACATGAGC	(AT) 21	205-217	2	205-217	2
	R: GGTGGAGTACAACCATCCTTTG					

La amplificación en *singleplex* se llevó a cabo en un volumen final de 15 µL que incluyó: 10.75 µL de agua Mili Q estéril, 1.5 µL de tampón (10X), 0.9 µL de MgCl₂ (25 mM), 0.75 µL de dNTPs (2 mM), 0.2 µL de cada cebador (50 µM), 0.2 µL de AmpliTaq Gold™ DNA Polymerase (5 u/µL, Applied Biosystems™) y 0.5 µL de muestra de ADN (40 ng/µL) de cada uno de los individuos.

La PCR se realizó en un termociclador Applied Biosystems™ 2720 Thermal Cycler utilizando el programa: 95°C 10 min; 35 ciclos de 94°C 45s, temperatura de hibridación de cada marcador (Tabla 3) 50s, 72°C 50s y finalmente 72°C 10 min. Como se muestra en la Tabla 3 para cada locus se testaron 2 temperaturas distintas, una la recomendada por García-Verdugo *et al.* (2015) y otra dos grados mayor o menor de la recomendada dependiendo del locus, nunca superando los 60°C. El objetivo era encontrar temperaturas comunes para poder agrupar el mayor número de loci de cara a realizar la amplificación mediante PCR *multiplex* (amplificación conjunta de varios loci microsatélite), optimizando así el coste de su amplificación.

Tabla 3. Temperaturas óptimas de hibridación testadas en este estudio vs. Recomendadas por García-Verdugo *et al.* (2019) para la amplificación cruzada de los marcadores microsatélites de *D. rodriguezii* en *D. gnidium* y las testadas en el presente trabajo.

Microsatélites	Temperatura de hibridación recomendada (°C)	Temperaturas testadas (°C)	
Dro-12	52,6	52	54
Dro-41	50	50	52
Dro-46	50,1	50	52
Dro-25	59	58	60
Dro-28	60	58	60
Dro-35	59	58	60
Dro-42	60	58	60
Dro-73	59	58	60
Dro-19	55,7	56	58

La evaluación de la amplificación se realizó migrando los fragmentos amplificados en una electroforesis en geles de agarosa al 1% con tampón TAE (0,5X) y con tinción SYBR™ safe (0,5X) para la visualización de ADN bajo una lámpara de luz azul. De cada PCR *singleplex* se comprobaron 7 µL mezclados con 1 µL de tampón de carga azul de bromofenol 6X. En uno de los pocillos del gel se incluyó un marcador de peso molecular, GeneRuler 100bp *DNA Ladder*, para estimar el tamaño de los fragmentos.

Una vez comprobada la correcta amplificación de los loci en PCR *singleplex* y establecida la temperatura óptima de hibridación, en los marcadores con amplificación positiva, se marcó uno de los cebadores de cada pareja con un fluoróforo para su análisis en un secuenciador automático. La elección del fluoróforo (4 disponibles: FAM, VIC, NED y PET) se realizó teniendo en cuenta la combinación de los distintos loci para su amplificación en PCR *multiplex*, atendiendo a su temperatura de hibridación y al rango de tamaño indicado en García-Verdugo *et al.* (2019), de manera que sea posible analizar más de cuatro loci en

cada *multiplex* al combinar marcaje (4 fluoróforos) y rango de tamaño de los alelos de cada loci (no solapamiento). De esta forma, se consigue optimizar el coste del proceso.

La amplificación utilizando los cebadores marcados fue testada tanto en PCR *singleplex* como en *multiplex* en 8 muestras de la población de Cíes. El procedimiento llevado a cabo para la amplificación *singleplex* fue el mismo que el descrito anteriormente para los cebadores sin marcar. El método para la amplificación en *multiplex* se describe a continuación.

Se desarrollaron dos PCR *multiplex* a 2 temperaturas distintas, *multiplex* 1 (M1) a 52°C y *multiplex* 2 (M2) a 58°C (Tabla 4), teniendo en cuenta los resultados obtenidos en la amplificación *singleplex*, con cebadores sin marcar, testando dos temperaturas para cada locus indicada anteriormente (Tabla 3).

Tabla 4. Condiciones de la amplificación en *multiplex*, indicando el marcaje fluorescente seleccionado para cada marcador y la temperatura de hibridación utilizada en cada *multiplex*.

<i>Multiplex</i>	Locus	Fluoróforo	T ^a (°C)
M1	Dro-12	Pet	52
	Dro-41	Vic	
	Dro-46	Fam	
M2	Dro-25	Fam	58
	Dro-28	Pet	
	Dro-35	Ned	
	Dro-42	Vic	
	Dro-73	Fam	

Las PCR *multiplex* se llevaron a cabo utilizando el kit GoTaq® G2 Flexi DNA Polymerase (Promega), siguiendo el protocolo del fabricante, en un volumen final de 10 µL que incluyó: 1 µL de una mezcla de los cebadores de cada *multiplex* a una concentración de 2 mM, 3.5 µL de agua Mili Q, 5 µL de Master Mix y 0,5 µL de ADN a una concentración de 40 ng/µL. La amplificación se realizó según las condiciones indicadas en el kit en un termociclador

Applied Biosystems™ 2720 Thermal Cycler: 95 °C 2 min; 94 °C 30 s, temperatura de hibridación (Tabla 4) 1 min y 30 s, 72°C 1 min, estos tres últimos pasos se repitieron 34 veces y, 72°C 5 min.

Las amplificaciones *multiplex* se comprobaron en geles de agarosa al 1% utilizando 5 µL de la PCR mezclados con 1 µL de tampón de carga azul de bromofenol 6X. Además, se valoró si era necesaria su dilución en agua mili Q estéril antes de su análisis en secuenciador automático para evitar la saturación de la señal fluorescente.

Tras la comprobación de las PCR *multiplex* se procedió a la preparación de las muestras para su genotipado mediante análisis de fragmentos de ADN en un secuenciador automático Applied Biosystems™ 3500xL en la Unidad de Secuenciación y Análisis de Fragmentos de la Universidad de Santiago de Compostela (Campus de Lugo). En cada pocillo de la placa del secuenciador se introdujo 0.6 µL del producto de la PCR *multiplex* de cada muestra, 8.75 µL de formamida y 0.25 µL de marcador interno de tamaño GeneScan 500 Liz Size Standard (Applied Biosystems™) utilizado para determinar el tamaño de los fragmentos amplificados, entre 35 y 500 pb, de los marcadores microsatélites a analizar (Tabla 4). Esta mezcla se sometió a una desnaturalización a 95°C 3 min en un termociclador Applied Biosystems™ 2720 Thermal Cycler con la tapa abierta. Posteriormente, las muestras se introdujeron en el secuenciador automático para su análisis. Los alelos para cada marcador se visualizaron en una gráfica llamada electroferograma, que muestra picos correspondientes a los alelos de los loci analizados.

En esta gráfica (Fig. 5) el eje de abscisas representa el tamaño en pb y el de ordenadas la intensidad de la fluorescencia (cantidad de amplificación). El genotipado se basa en correlacionar el pico del alelo del microsatélite con los del marcador (*sizer*) utilizado, determinando así el tamaño en pb de los alelos. Mediante el programa GeneMapper® Software 6 se creó un panel de análisis con los datos obtenidos, se genotiparon los individuos de forma automática y, finalmente se revisaron de forma manual previamente a los análisis posteriores.

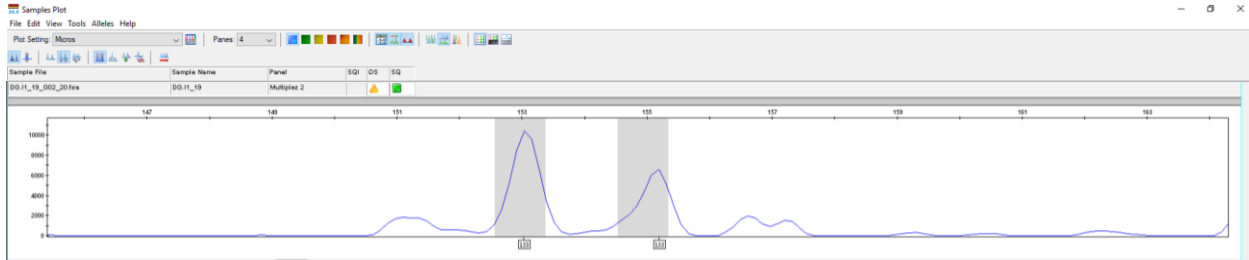


Figura 5. Ejemplo de electroferograma obtenido tras el genotipado en secuenciador automático Applied Biosystems™ 3500xL en la Unidad de Secuenciación y Análisis de Fragmentos de la Universidad de Santiago de Compostela (Campus de Lugo).

3.4 Caracterización poblacional de marcadores genéticos

En este estudio se emplearon dos tipos de marcadores moleculares: microsatélites y SNPs. Se analizó el genotipado de los loci microsatélites obtenidos en las 6 localidades estudiadas (ver Resultados), (Tabla 1). Además, se analizaron los datos de 2011 SNPs genotipados en 4 de las localidades muestreadas (Tabla 1).

Para los microsatélites analizados en las localidades estudiadas, los posibles errores de genotipado (tartamudeo, alelos fuera de rango) fueron analizados y revisados manualmente, y la posible presencia de alelos nulos inferida a partir del análisis poblacional (defecto de heterocigotos).

Los datos genotípicos de SNPs analizados en este estudio proceden de un estudio paralelo realizado mediante genotipado por secuenciación con tecnología 2bRADseq (Wang *et al.*, 2012) en el marco del proyecto LIFE INSULAR (datos no publicados). El panel de SNPs necesario para el estudio se obtuvo del procesado bioinformático de los datos de secuenciación masiva de las 64 muestras de *D. gnidium* analizadas (archivos FASTQ) obtenidos con tecnología Illumina NextSeq, en la plataforma FISABIO (<https://fisabio.san.gva.es/es/>).

La caracterización de SNPs en este estudio se inició con el panel filtrado de SNPs en diferentes formatos de archivo para los análisis subsiguientes de calidad y diversidad genética (VCF, GENEPOP, STRUCTURE). Se utilizó el programa VCFtools - 0.1.16 (Danecek

et al., 2011) para evaluar la calidad del panel de SNPs, analizando el porcentaje de datos perdidos (*missing data*) y la ratio de transiciones/transversiones (Ts/Tv).

3.4.1 Análisis de diversidad genética

A partir de los datos genotípicos de los marcadores, tanto microsatélites como SNPs, se realizaron los mismos análisis de diversidad, diferenciación y estructuración genética poblacional, pero de manera independiente.

Las frecuencias alélicas por locus fueron calculadas a partir de las frecuencias genotípicas observadas. Para microsatélites se utilizó el programa FSTAT v.2.9.4 (Goudet, 2003). Con respecto a los SNPs se calcularon utilizando la función "basic_info" del paquete de R genepop v.1.2.2 (Rousset *et al.*, 2008).

Se evaluó el ajuste al equilibrio Hardy-Weinberg (HW), según el cual en una población de organismos diploides y de tamaño infinito supone apareamiento al azar (*panmixia*), sometida únicamente a las leyes de transmisión de Mendel, en las que se mantienen constantes las frecuencias génicas y genotípicas. Para un locus con dos alelos A y B, tras un apareamiento al azar, con frecuencias alélicas p y q, las frecuencias esperadas bajo equilibrio HW de los genotipos AA, AB y BB serán p^2 , $2pq$, q^2 respectivamente (Pierce *et al.*, 2021). Para los loci microsatélite, la probabilidad de ajuste al equilibrio HW por locus y localidad, así como global sobre loci por población (test de Fisher), se analizó mediante el programa GENEPOP 4.7 (versión para sistema operativo Windows/online: <https://genepop.curtin.edu.au/>) . La magnitud de las desviaciones de equilibrio HW (F_{IS} o coeficiente de endogamia) por locus y localidad para los loci microsatélite se estimó usando el programa GENEPOP 4.7. Para los marcadores SNPs se evaluó la probabilidad global de ajuste al equilibrio HW, así como el valor de F_{IS} con sus intervalos de confianza al 95% utilizando los paquetes de R genepop v.1.2.2 y DiveRsity v.1.9.90 (Keenan *et al.*, 2013) en RStudio.

Los estimadores de diversidad genética (número medio de alelos por locus (N_a), heterocigosis observada (H_o) y esperada (H_e)) por muestra poblacional se obtuvieron

usando el programa FSTAT v.2.9.4 y DiveRsity v. 1.9.90 para microsatélites y SNPs, respectivamente.

Debido a las diferencias en el tamaño muestral causadas por los datos ausentes (*missing data*) identificados en los marcadores SNPs, se calculó el estimador de riqueza alélica (A_r) empleando el método de rarefacción recomendable para tamaños de muestra desiguales utilizando el paquete DiveRsity v.1.9.90. Esta metodología no se ha empleado con microsatélites, al no detectarse datos ausentes ni haber diferencias en tamaño muestral en las muestras estudiadas. La riqueza alélica es un estimador de la diversidad alélica por locus, corregida por el tamaño muestral mínimo disponible entre las poblaciones analizadas cuando se emplea el método de rarefacción. Se trata de una medida más informativa de la diversidad que el estimador N_a , cuando existen diferencias en tamaño muestral entre poblaciones.

3.4.2 Análisis de diferenciación y estructura genética

Se analizó el coeficiente de diferenciación interpoblacional (F_{ST}) a nivel global y entre todos los pares de localidades estudiadas utilizando, en el caso de los loci microsatélites, el programa FSTAT v.2.9.4 a través del índice de fijación de Weir y Cockerham (W&C) con una significación del 5% ($p < 0,05$) usando 1000 permutaciones. Por otra parte, para los marcadores SNPs se utilizó el programa Genepop para calcular el valor de la F_{ST} global y el programa StAMPP v.1.6.3 (Pembleton *et al.*, 2013) para calcular los valores de F_{ST} por pares de localidades (Weir & Cockerham, 1984), usando 10 000 *bootstraps* para obtener la significación con intervalos de confianza al 95%. Para ello, se convirtió el archivo GENEPOP de entrada en objeto de R de clase *genind* y finalmente a clase *genlight* utilizando las librerías dartR v.2.9.7 (Gruber *et al.*, 2018) y adegenet v.2.1.10 (Jombart, 2008).

El análisis de estructuración genética se realizó únicamente con los marcadores SNPs, descartando su aplicación con microsatélites, dado el reducido número de loci informativos detectados en este estudio. Para determinar el número óptimo de unidades poblacionales genéticas (K) Se utilizó el método de agrupamiento bayesiano del programa

STRUCTURE v.2.3.4 (Pritchard *et al.*, 2000; Falush *et al.*, 2007) implementado en el CESGA (Centro de Supercomputación de Galicia) con el paquete de R ParallelStructure v.1.0.0 (Besnier & Glover, 2013) que permite paralelizar este software empleando múltiples procesadores para agilizar los análisis de datos genómicos. Se testó un rango de valores de K de 1 a 7 (Número de localidades + 3) con un *burn-in* de 100 000 interacciones, cadena de Markov posterior de 200 000 interacciones y 10 réplicas por cada valor de K, usando como modelo de ancestralidad ADMIXTURE y como modelo de frecuencias alélicas CORRELATED.

Los datos obtenidos con STRUCTURE se analizaron utilizando la herramienta en línea StructureSelector (<https://lmme.ac.cn/StructureSelector/>; Li & Liu, 2018). Esta herramienta permite evaluar el valor de K más probable trazando el gráfico de Probabilidades ($\ln(PD)$) y calculando el estadístico ΔK , propuesto por Evanno *et al.* (2005). El valor de K que maximiza el valor de ΔK se considera como el número óptimo de grupos genéticos presentes en la muestra. A su vez, con esta herramienta en línea se obtuvieron los gráficos correspondientes reflejando la pertenencia más probable de cada individuo a los diferentes grupos genéticos (CLUMPAK; Kopelman *et al.*, 2015).

Adicionalmente, se realizó un análisis multivariante discriminante de componentes principales (DAPC) usando el paquete de R ADEGENET v.2.1.10. Los análisis DAPC se realizaron utilizando las localidades de muestreo como grupos a priori. Se utilizó la función "find.clusters" para determinar el número de grupos genéticos (K), seleccionando como K óptimo aquel con el valor BIC (criterio de información bayesiano) más bajo. El número óptimo de PCs (Componentes Principales) a utilizar en el DAPC se determinó de modo que la proporción de varianza conservada fuera superior al 50%. Los gráficos resultantes se representaron en un diagrama de dispersión en dos dimensiones utilizando los mejores *discriminant eigenvalues* (componentes discriminantes) del DAPC.

4. RESULTADOS Y DISCUSIÓN

4.1 Extracción de ADN

Los resultados obtenidos en la extracción de ADN a partir de hojas desecadas mediante el kit E.Z.N.A.® Plant DNA DS Kit (Omega), con las modificaciones introducidas para mejorar el rendimiento y calidad del ADN, mostraron que el protocolo aplicado permite obtener ADN de calidad, pureza y en cantidad suficiente para ser utilizado en el posterior análisis de marcadores moleculares.

4.2 Amplificación y genotipado de marcadores microsatélite ortólogos

En ausencia de microsatélites específicos en *D. gnidium* en este estudio se utilizaron 9 marcadores microsatélite de los 15 descritos en la especie congénérica *D. rodriguezii* por García-Verdugo *et al.* (2019) y testada su amplificación cruzada en *D. gnidium*. En la puesta a punto inicial, utilizando 3 muestras de la población de Cíes y PCR *singleplex* con cebadores sin marcar, testando dos temperaturas de hibridación para cada locus, se observó amplificación positiva en geles de agarosa en 8 de los 9 loci (88,9%).

Como se muestra en la Figura 6, para cada uno de estos 8 loci se obtuvo un producto único de amplificación en el tamaño esperado de acuerdo con los resultados obtenidos por García-Verdugo *et al.* (2019). El marcador Dro-19, sin embargo, mostró amplificación inespecífica y fuera del tamaño descrito por García-Verdugo *et al.* (2019) en las dos temperaturas testadas, por lo que se descartó para el análisis poblacional. Por lo tanto, los resultados del presente estudio en geles de agarosa confirman los resultados obtenidos por García-Verdugo *et al.* (2019) salvo en el locus Dro-19.

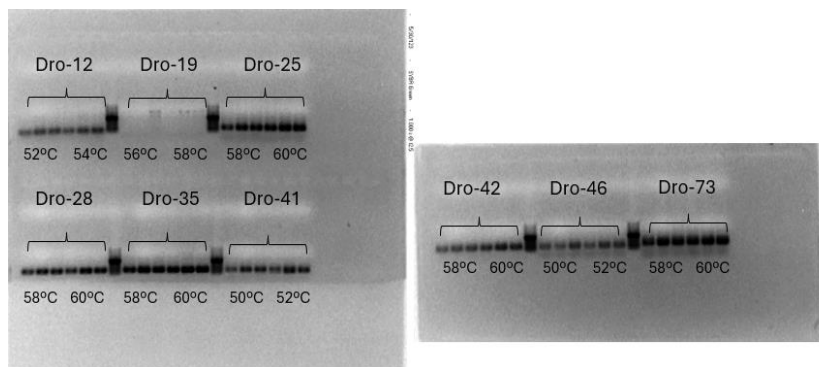


Figura 6. Visualización de los fragmentos amplificados mediante PCR *singleplex* en geles de agarosa al 1 % de los marcadores seleccionados utilizando cebadores sin marcar y testando 2 temperaturas de hibridación para cada microsatélite.

A continuación, una vez seleccionada la temperatura óptima de hibridación en cada uno de los 8 marcadores, se testó su amplificación en 8 muestras de la misma población utilizando uno de los cebadores marcado con un fluoróforo, necesario para su genotipado en un secuenciador automático. En primer lugar, se comprobó en geles de agarosa la correcta amplificación en PCR *singleplex* con el cebador marcado. A continuación, se diseñaron 2 PCR *multiplex* agrupando los loci microsatélite teniendo en cuenta la temperatura óptima de hibridación, el tamaño del producto de amplificación y el marcaje, tal como se indicó en apartado de Material y Métodos (Tabla 3). Las amplificaciones *multiplex* se comprobaron en geles de agarosa y se procedió a su análisis en un secuenciador automático Applied Biosystems™ 3500xL, junto con las PCR *singleplex* realizadas con cebador marcado, para evaluar el polimorfismo. En el caso de las PCR *singleplex* estas se agruparon para su análisis en el secuenciador de acuerdo con las *multiplex* diseñadas.

Los resultados observados en el secuenciador automático, en las 8 muestras analizadas, mostraron que de los 8 marcadores microsatélites, con amplificación positiva en geles de agarosa, 7 presentaron picos correspondientes a alelos en el tamaño o próximo al esperado para cada marcador de acuerdo con García-Verdugo *et al.* (2019), mientras que el marcador Dro-46 presentó múltiples picos y poco específicos, por lo que se descartó del análisis poblacional. Sin embargo, como se muestra en la Tabla 5, de estos 7 marcadores 5 resultaron monomórficos, incluso en un análisis posterior utilizando 3 muestras de cada

una de las 6 poblaciones, no resultando útiles para el análisis poblacional objetivo de este trabajo al no presentar variación. Solo 2 marcadores (28,6%), Dro-25 y Dro-42, resultaron polimórficos (Fig. 7) y se aplicaron en los análisis poblacionales.

Tabla 5. Comparación del tamaño alélico, expresado en bp, obtenido en el trabajo de García-Verdugo *et al.* (2019) y en el presente trabajo para los 8 marcadores con amplificación positiva y loci que resultaron polimórficos (*). (-: marcador descartado del análisis poblacional).

Loci	García-Verdugo <i>et al.</i> (2019)	Presente trabajo
Dro-12	255	269
Dro-41	184	184
Dro-46	173, 179	-
Dro-25	152, 156, 168	153, 155*
Dro-28	198	199
Dro-35	192	192
Dro-42	164	156, 165, 169*
Dro-73	205, 217	205, 217

En resumen, en este trabajo, de los 9 marcadores seleccionados, descritos y aplicados en *D. rodriguezii*, y testada su amplificación cruzada en *D. gnidium* por García-Verdugo *et al.* (2019), 8 amplificaron correctamente y solo 2 resultaron finalmente polimórficos y útiles para el análisis poblacional. Los resultados obtenidos en los análisis de amplificación cruzada concuerdan, salvo en un marcador, con los obtenidos por García-Verdugo *et al.* (2019), confirmando una notable conservación de las regiones flanqueantes de los loci microsatélites entre las dos especies congénicas. Por otro lado, el grado de polimorfismo en *D. gnidium* fue muy inferior al observado en *D. rodriguezii*. Estas diferencias no son raras dado que la amplificación cruzada no conlleva necesariamente una conservación del grado de polimorfismo por locus, siendo característico de cada especie como se ha observado en numerosas especies (Ellegren *et al.*, 1995; Crawford *et al.*, 1998). Por tanto, a pesar de que en el trabajo de García-Verdugo *et al.* (2019) se sugiere la utilidad de estos marcadores,

dada la tasa de amplificación cruzada obtenida, en estudios poblacionales en especies del género *Daphne*, en este trabajo se concluye que su utilidad es limitada en *D. gnidium*.



Figura 7. Electroferogramas obtenidos en el secuenciador automático Applied Biosystems™ 3500xL correspondientes a los loci polimórficos Dro-25 y Dro-42 en algunas muestras analizadas. Se indican los tamaños, en pb, de los alelos (picos) observados. Se muestran los perfiles obtenidos para los homocigotos de cada uno de los alelos de cada marcador y los heterocigotos. A: Perfiles alélicos del locus Dro-25. B: Perfiles alélicos del locus Dro-42.

4.3 Análisis de resultados de marcadores microsatélites

4.3.1 Frecuencias alélicas

A partir de los datos genotípicos de los loci microsatélites polimórficos en las 120 muestras analizadas (Dro-25 y Dro-42), se calcularon las frecuencias alélicas de ambos loci en cada una de las seis localidades. En la Tabla 6 se muestran las frecuencias alélicas para ambos loci.

Tabla 6. Frecuencias alélicas por locus y muestra poblacional analizada de *D. gnidium*. Para cada locus se indican los alelos detectados (p: alelo detectado).

	Localidades					
	C1	C2	C3	I1	I2	I3
	Dro-25					
p: 153	0,500	0,625	0,525	0,550	0,525	0,475
p: 155	0,500	0,375	0,475	0,450	0,475	0,525
	Dro-42					
p: 156	0,075	0,025	0,050	0,225	0,025	0,000
p: 165	0,825	0,975	0,950	0,650	0,975	1,000
p: 169	0,100	0,000	0,000	0,125	0,000	0,000

Para el locus Dro-25 se observaron 2 alelos, el alelo 153 con frecuencias intermedias (próximas a 0,5), bastante homogéneas entre localidades, y valor máximo en C2, siendo el alelo 155 el menos frecuente.

En el caso del locus Dro-42 se observaron 3 alelos en total, con un alelo 165 preferente en todas las localidades, fijado en I3, mientras que los dos alelos restantes mostraron distribución diferencial entre las muestras estudiadas. Los 3 alelos se detectaron tanto en C1 como en I1.

El análisis de las frecuencias alélicas (Tabla 6) mostró poca variación alélica para ambos loci en todas las localidades, como describieron en las poblaciones de la especie focal de aislamiento de dichos marcadores analizadas por García-Verdugo *et al.* (2019). Las diferencias de las frecuencias alélicas entre poblaciones pueden contribuir a variaciones en la diversidad genética entre las localizaciones insulares (García-Verdugo *et al.*, 2019). En

las frecuencias alélicas fueron bastante homogéneas entre localidades para el locus Dro-25, con mayor heterogeneidad para Dro-42, con similitud entre Cíes y Barra (I1-C2), diferenciadas del resto, especialmente en la isla de Sálvora (I3).

4.3.2 Análisis de diversidad genética

A partir de los genotipos de los microsatélites polimórficos disponibles se analizaron por primera vez distintos estimadores de diversidad genética por locus y localidad muestreada de la especie *D. gnidum* (Tabla 7).

Para el número medio de alelos por locus (N_a) se observó cierta heterogeneidad entre loci. Para el locus Dro-25, N_a se mantuvo constante con 2 alelos en todas las muestras estudiadas mientras que para el locus Dro-42 N_a osciló entre 1 en la muestra I3 de Sálvora y 3 en C1 de Barra

Los valores de heterocigosis observada (h_o) para el locus Dro-25 variaron de 0,400 en C1 a 0,850 en I2 e I3, y para el locus Dro-42 oscilaron entre 0 en I3 y 0,550 en I1, mientras que la heterocigosis esperada (h_e) para Dro-25 oscilaron entre 0,476 en C2 y 0,516 en C1, y para Dro-42 entre 0 en I3 a 0,524 en I1 (Tabla 7).

Los valores de heterocigosis esperada para el locus Dro-25 fueron similares a los descritos en García-Verdugo *et al.* (2019). No ocurre lo mismo para el locus Dro-42. Esto puede deberse a que se trata de una comparación de especies distintas en ambientes diferentes, que pueden estar sujetas a factores evolutivos específicos, por lo que no tendrían por qué ser similares.

En la Tabla 7 también se muestran los resultados de los tests de ajuste de frecuencias genotípicas observadas respecto a las esperadas en el equilibrio HW por locus y localidad en los dos loci microsatélites polimórficos. Para el locus Dro-42 no se detectaron desviaciones significativas de HW ($p > 0,05$). Solo se detectaron dos tests significativos al 5% para el locus Dro-25 en las localidades insulares I2 e I3, ambos significativos tras la corrección de Bonferroni. Los valores F_{IS} por locus fueron negativos indicando en ambos

casos exceso de heterocigotos (-0,691). Esta desviación puede deberse a la varianza de muestreo por tamaño muestral limitado, sin descartar explicaciones alternativas en ambas localidades insulares para este locus, incluyendo efectos específicos de distintas fuerzas evolutivas (flujo génico, selección a favor de heterocigotos y/o apareamiento clasificado), o bien asociadas con baja tasa de autofecundación, tal como se ha descrito en la especie (Roccoliello *et al.*, 2009).

Tabla 7. Análisis de diversidad genética por locus y localidad para los dos loci microsatélite polimórficos analizados de manera independiente en muestras de *D. gnidium*. Para cada locus y localidad se indica: el número medio de alelos (N_a), la heterocigosis observada (h_o), la heterocigosis esperada (h_e), la probabilidad (P_{HW}) de significación del test de ajuste de las frecuencias genotípicas observadas a las esperadas en el equilibrio HW, así como el valor de F_{IS} por locus (-: no se determinó el test global de ajuste al equilibrio HW).

Localidades	Dro-25					Dro-42				
	N_a	h_o	h_e	P_{HW}	F_{IS}	N_a	h_o	h_e	P_{HW}	F_{IS}
C1	2	0,400	0,516	0,394	0,225	3	0,200	0,315	0,071	0,364
C2	2	0,650	0,476	0,158	-0,365	2	0,050	0,050	-	-
C3	2	0,650	0,508	0,367	-0,280	2	0,100	0,097	1,000	-0,028
I1	2	0,500	0,508	1,000	0,016	2	0,550	0,524	1,000	-0,057
I2	2	0,850	0,503	0,005	-0,691	2	0,050	0,050	-	-
I3	2	0,850	0,503	0,005	-0,691	1	0,000	0,000	-	-

En la Tabla 8 se muestran los valores promedios sobre loci de diversidad genética en las distintas localidades muestrales analizadas. El número medio de alelos (N_a) osciló entre 2,5 en la muestra C1 de Barra y 1,50 en I3 de Sálvora, sugiriendo menor diversidad en un bajo número de alelos para la realización de un análisis poblacional.

Los valores de heterocigosis observada (H_o) y esperada (H_e) por localidad fueron similares, concordantes con el test global de ajuste al equilibrio HW para todas las muestras continentales e I1 en Cíes. Solo se detectó desviación significativa global de HW en las dos localidades insulares más al norte (Ons y Sálvora). En ambos casos, se confirman los valores negativos de F_{IS} , indicando un exceso de heterocigotos asociado al locus Dro-25, coincidiendo con el análisis por locus en las mismas muestras, I2 e I3 (Tabla 7). Dado el bajo número de loci informativos disponibles no es posible descartar un efecto de varianza muestral en estos resultados para el locus Dro-25, aunque la desviación de HW lo convierte en un marcador sugerente bajo influencia de fuerzas evolutivas para estudios futuros en distintas poblaciones de *D. gnidium*.

Tabla 8. Valores medios de diversidad genética por localidad para loci microsatélites polimórficos analizados en muestras de *D. gnidium*. Se presentan el número de alelos por locus (N_a); heterocigosis observada (H_o), y heterocigosis esperada (H_e), indicando entre paréntesis la desviación estándar (SD). Además, se indica la probabilidad global sobre loci de ajuste al equilibrio HW (P_{HW}), y el valor F_{IS} promedio sobre loci (-: no calculado, solo un locus informativo).

Localidad	N_a (SD)	H_o (SD)	H_e (SD)	P_{HW} Global	F_{IS} (SD)
C1	2,50 (0,71)	0,300 (0,141)	0,415 (0,143)	0,233	0,294 (0,099)
C2	2,00 (0,000)	0,050 (0,424)	0,263 (0,301)	0,158	-0,365 (-)
C3	2,00 (0,000)	0,375 (0,389)	0,303 (0,291)	0,683	-0,154 (0,178)
I1	2,00 (0,000)	0,525 (0,035)	0,516 (0,011)	1	-0,021 (0,051)
I2	2,00 (0,000)	0,450 (0,566)	0,277 (0,320)	0,005	-0,691 (-)
I3	1,50 (0,71)	0,425 (0,601)	0,251 (0,355)	0,005	-0,691 (-)

Los valores promedio del estimador de diversidad genética (H_e) mostraron valores superiores en la muestra C1 en Barra seguida de I1 en Cíes (Tabla 8; $H_e > 0,4$), respecto al resto de localidades analizadas, tanto continentales como insulares (H_e entre 0,251 y 0,303; Tabla 8). En comparación con los valores censales de IBADER (datos no publicados)

Barra y Cíes presentan valores poblacionales más altos, mientras que Ons y Sálvora presentan valores más reducidos. Los resultados con microsatélites en este TFG son pioneros en *D. gnidium* en su rango de distribución natural, al no disponer de estudios poblacionales previos con microsatélites en esta especie. Estos resultados sobre loci, se encuentran en el rango próximo a los promedios poblacionales con microsatélites en otras especies vegetales en el Noroeste ibérico, de interés medicinal, como *Arnica montana* L. con tamaños poblacionales reducidos (Bouza *et al.*, 2023). Sin embargo, las comparaciones entre especies distintas deben interpretarse con cautela al estar sujetas a procesos evolutivos específicos de cada linaje.

En la Tabla 9 se muestran los valores promedio poblacionales del número medio de alelos (Na) y heterocigosis esperada (He), donde se obtuvieron ligeras diferencias entre los grupos poblacionales continentales vs. insulares, especialmente en diversidad alélica (Na 2,167 vs. 1,833; He 0,327 vs. 0,348; Tabla 9).

Con los valores medios obtenidos por grupo, se deduce la existencia de menor diversidad asociada al Na y He en las poblaciones insulares, especialmente en I2 (Ons) e I3 (Sálvora). La falta de diversidad puede estar relacionado con la ausencia o disminución de los alelos poco frecuentes afectando a la capacidad de la población para adaptarse a nuevas condiciones ambientales y evolutivas (García-Verdugo *et al.*, 2015; Vera *et al.*, 2020; El Mouzazi *et al.*, 2023). En el par C3-I3 (Corrubedo- Sálvora) se observó una menor diversidad alélica, indicando menor variabilidad genética y la presencia de alelos poco frecuentes en comparación con el resto de localidades.

En García-Verdugo *et al.*, (2019) los resultados medios de diversidad obtenidos en las dos poblaciones analizadas de *D. rodriguezii* (A= 22 individuos y B= 20 individuos) fueron mayores a los obtenidos en este TFG utilizando los locus comunes, Dro-25 y Dro-42, especialmente para el número de alelos (Na poblaciones A y B vs. media en este estudio: 6,5 y 5,5 vs. 2; He: 0,437 y 0,411 vs. 0,338). Esto se puede explicar por el sesgo de selección de alto polimorfismo en la especie focal *D. rodriguezii* que puede no mantenerse en otras

especies. En todo caso, los resultados de García-Verdugo *et al.*, (2019) sí demostraron que, a mayor tamaño de población, se encontraba mayor diversidad genética, encajando con los resultados de este estudio en *D. gnidium*, los primeros disponibles en esta especie. No obstante, estas comparaciones se deben tomar con cautela, al referirse a especies distintas. Otros factores a tener en cuenta incluyen el estado del endemismo, el rango de distribución insular y los procesos históricos, incluyendo la perturbación provocada por el ser humano, los cuales parecen mostrar una mayor influencia en los niveles de diversidad genética más allá del tamaño poblacional y la condición insular (García-Verdugo *et al.*, 2015). En este estudio, hay que destacar que las poblaciones insulares analizadas se encuentran en áreas protegidas en el Parque Nacional de las Islas Atlánticas, con impacto antropológico controlado, aunque no se puede descartar influencia potencial del turismo ecológico.

Tabla 9. Análisis de diversidad genética media por grupos poblacionales para los dos microsatélites polimórficos en muestras de *D. gnidium*. Se agrupan todas las muestras insulares y continentales, así como por pares adyacentes de continentales e insulares.

Localidades	Na	He
Continental	2,167	0,327
Insulares	1,833	0,348
Par C1-I1	2,250	0,466
Par C2-I2	2,000	0,270
Par C3-I3	1,750	0,277

El análisis de estimadores de diversidad agrupando pares de muestras insulares y continentales adyacentes, mostró mayor diversidad genética en el par C1-I1 correspondiente a Barra- Cíes, que puede ser debida a su proximidad biogeográfica y conectividad histórica reciente en la última glaciación (Arce Chamorro *et al.*, 2021). Esto habría facilitado un mayor flujo genético entre ambas poblaciones, traduciéndose en una mayor variabilidad genética. Por el contrario, una mayor distancia geográfica entre el resto de pares de localidades o una reciente colonización de esas áreas pueden determinar flujo

génico restringido y menor diversidad genética como se ha descrito en otros escenarios insulares en plantas (García-Verdugo *et al.*, 2015).

El mantenimiento de la diversidad genética puede representar un desafío para las poblaciones ubicadas en regiones aisladas y en límites del área de distribución o de la especie como es el caso de los territorios insulares donde los eventos de aislamiento y extinción/colonización son frecuentes. Sin embargo, este tipo de poblaciones con cierto grado de aislamiento local o regional pueden preservar variantes genéticas específicas de alto interés para su conservación, como se ha descrito en otras plantas medicinales (Bouza *et al.*, 2023).

El uso de marcadores moleculares permite evaluar la distribución de la variabilidad genética entre las poblaciones naturales. Los loci microsatélites proporcionan información sobre procesos demográficos recientes como el flujo y la deriva genética, pero también pueden aportar huellas de selección divergente con impacto en la estimación de diversidad y diferenciación genética (García-Verdugo *et al.*, 2015; Bouza *et al.*, 2023).

Las diferencias en patrones de diversidad genética entre las distintas localidades en este estudio, especialmente para el locus Dro-42, se complementaron con la evaluación del grado de diferenciación genética interpoblacional en el área de estudio.

4.3.3 Análisis de diferenciación y estructura genética

El coeficiente de diferenciación interpoblacional (F_{ST}) obtenido a nivel global en este estudio fue de 0,024, revelando que el 2,4% de la diversidad total a nivel global se debe a diferencias entre poblaciones estudiadas. Esto refleja un nivel bajo de diferenciación genética entre las localidades analizadas.

Los valores de F_{ST} por pares entre todas las muestras poblacionales oscilaron entre 0 (asumiendo los valores negativos como no significativamente distintos de 0) y 0,082 (Tabla 10). Los valores de F_{ST} fueron moderados y significativos para los pares I1-C2, I1-C3, I1-I2 y I1-I3 ($p < 0,05$). En estos cuatro casos pareados el resultado obtenido mostró valores de F_{ST}

inferiores a 0,05, que indican una divergencia genética baja entre dichas poblaciones y flujo genético entre las diferentes muestras poblacionales (Vilas & Álvarez., 2023).

Tabla 10. Análisis de F_{ST} por pares. En negrita se indican los resultados significativos ($p < 0,05$).

Localidades	C1	C2	C3	I1	I2	I3
C1	-	NS	NS	NS	NS	NS
C2	0,022	-	NS	**	NS	NS
C3	-0,007	0,001	-	*	NS	NS
I1	0,004	0,082	0,058	-	**	**
I2	0,003	0,006	-0,014	0,077	-	NS
I3	0,013	0,032	-0,005	0,100	-0,003	-

*** p-value < 0.001, ** p-value < 0.01, * p-value < 0.05, NS: No significativo.

El análisis global de F_{ST} por locus mostró una contribución preferente del marcador Dro-42 (F_{ST} : 0,103) a la diferenciación interpoblacional, concordante con la distribución diferencial de frecuencias alélicas entre localidades para este locus (Tabla 6), frente a la homogeneidad observada para Dro-25 ($F_{ST} \sim 0$). Estos resultados podrían estar relacionados con patrones evolutivos específicos de cada marcador, pero dado el bajo número de loci, no se puede conocer su representatividad en el genoma de las poblaciones estudiadas. Para obtener resultados más consistentes se hace necesario ampliar el número de marcadores genéticos para evaluar el patrón de divergencia genética global. Entre las distintas opciones posibles, se podría identificar baja diferenciación y alta conectividad entre poblaciones con loci atípicos (Dro-42) sugiriendo selección divergente, así como, alta divergencia con flujo génico limitado entre las poblaciones estudiadas con loci atípicos (Dro-25) con patrón homogéneo entre localidades, sugiriendo selección equilibradora.

En García-Verdugo *et al.*, (2019) no se describen valores de F_{ST} entre poblaciones, impidiendo hacer una comparativa de diferenciación genética con poblaciones insulares de una especie congénica, que sería interesante en ausencia de datos microsatélites previos para *D. gnidium*. La comparación con valores de diferenciación descritos con este tipo de marcadores en otras plantas medicinales, como *A. montana*, revelan mucha mayor

diferenciación genética interpoblacional, asociada con presencia de poblaciones muy divergentes de esta especie en Galicia (Bouza *et al.*, 2023).

4.4 Análisis poblacional con marcadores SNPs

Debido al reducido número de loci microsatélites polimórficos en este estudio, se realizaron análisis adicionales utilizando un conjunto numeroso de marcadores SNP, siendo una aproximación sugerida para refinar los datos genéticos poblacionales en distintas especies, incluyendo otras plantas medicinales (Bouza *et al.*, 2023).

4.4.1 Análisis de diversidad genética

A partir del análisis de los genotipos de los 2011 SNPs, se caracterizó la calidad del panel final de loci, evaluando el porcentaje de *missing data* y la ratio de transiciones/transversiones (Tabla 11). Del análisis de los 2011 SNPs se determinó el número total de SNPs polimórficos por localidad, de utilidad en análisis de diversidad genética y estructura poblacional. Además, se analizó la ratio de transiciones/transversiones (Ts/Tv) mostró un resultado de $1400/611 = 2,291$ (Tabla 11). A pesar de que existe una mayor probabilidad de que ocurra una transversión que una transición, las transiciones suelen ser más comunes debido a la similitud estructural de las bases nitrogenadas implicadas.

Tabla 11. Análisis de las mutaciones que tienen lugar entre las diferentes muestras poblacionales: la primera y segunda letra se corresponde con las bases nitrogenadas de las variantes alternativas por SNP. Se muestra el número total de transiciones (Ts) y de transversiones (Tv).

Tipo de mutación	AC	AG	AT	CG	CT	GT	Ts	Tv
Número	161	715	131	157	685	162	1400	611

Por primera vez, en este estudio se analizaron diversos estimadores de diversidad genética de cuatro localidades muestreadas a partir de un panel muy numeroso de marcadores SNPs disponibles en *D. gnidium*. La proporción de SNPs polimórficos detectada por localidad fue muy alta en las cuatro muestras poblacionales estudiadas, máxima en C1 en

Barra (86,5%) y el valor más reducido en Sálvora (I3; 77,7%) (Tabla 12). Esto sugiere una menor diversidad genética en dicha localidad, posiblemente relacionada con factores ambientales, biogeográficos o demográficos.

La variabilidad genética observada con SNPs mostró una alta calidad de genotipado, con una baja proporción de datos ausentes (*missing data*), con un promedio de 5,2% y una desviación estándar del 8,1%. El SNP con mayor porcentaje de *missing data* fue del 17,2%, que representa un valor reducido en el contexto de estudios genético-poblacionales previos con paneles masivos de SNPs indicando que los datos son robustos para el análisis poblacional (Casanova *et al.*, 2021).

Los valores promedio de la heterocigosis observada (H_o) y esperada (H_e) al igual que en el análisis de microsatélites, fueron similares entre las muestras poblacionales, concordantes con el test global de ajuste de frecuencias genotípicas observadas a las esperadas en el equilibrio HW. Sin embargo, en las cuatro muestras se detectaron valores negativos de F_{IS} , que revelan una cierta tendencia de exceso de heterocigotos coincidiendo con el análisis de microsatélites en las muestras significativas de Ons y Sálvora (Tabla 8). Esto podría estar relacionado con sesgo dado el tamaño muestral disponible, pero también apuntar razones biológicas de desviación de apareamiento aleatorio relacionado con la estrategia reproductiva de esta especie con tasa reducida de autofecundación en flores hermafroditas, promoviendo la fecundación cruzada (Roccoliello *et al.*, 2009). Esta estrategia incrementaría la diversidad genética intrapoblacional evitando la autogamia (sin autofecundación por autoincompatibilidad) tal y como se describe en Gajdošová *et al.*, (2024), un estudio relacionado con la reproducción sexual de la especie congénica *Daphne arbuscula*.

Los valores promedio obtenidos por muestra poblacional del estimador de riqueza alélica mostraron valores similares en C1, I1 e I2, con el menor valor detectado en I3 (Sálvora) (Tabla 12), concordante con la proporción de loci polimórficos por localidad. Para el estimador de diversidad genética (H_e), los valores medios poblacionales fueron superiores

en la muestra I1 en Cíes, seguida de C1 en Barra (Tabla 12; $H_e > 0,25$), coincidiendo con los datos del análisis de microsatélites. Los resultados son pioneros en *D. gnidium* en su rango de distribución natural, al no estar descritos estudios poblacionales previos con SNPs en esta especie.

Tabla 12. Estimadores de diversidad genética para cada una de las muestras poblacionales: número de SNPs polimórficos (*: proporción relativa), riqueza alélica (A_r), heterocigosis observada (H_o) y esperada (H_e) así como las desviaciones del equilibrio de HW con su intervalo de confianza (IC) al 95%. Entre paréntesis se indica la desviación estándar (SD) de cada uno de los estimadores.

Localidades	SNPs polimórficos (*)	A_r (SD)	H_o (SD)	H_e (SD)	F_{IS} (IC)
C1	1740 (0,865)	1,865 (0,342)	0,365 (0,332)	0,251 (0,174)	-0,092 (-0,178; -0,021)
I1	1727 (0,859)	1,859 (0,348)	0,220 (0,311)	0,256 (0,178)	-0,165 (-0,240; -0,081)
I2	1627 (0,809)	1,809 (0,393)	0,170 (0,042)	0,204 (0,174)	-0,087 (-0,172; -0,005)
I3	1554 (0,777)	1,773 (0,419)	0,255 (0,078)	0,230 (0,180)	-0,151 (-0,250; -0,062)

Los valores obtenidos de heterocigosis esperada (H_e) en este estudio están dentro del rango descrito con paneles de SNPs en estudios poblacionales recientes poblacionales de especies relacionadas de la misma familia (*Daphnimorpha spp.*; Thymelaeaceae; Takahashi & Suyama., 2023), con aplicaciones en demografía y conservación.

Por otra parte, valores bajos de diversidad genética en especies vegetales de baja altitud pueden explicarse por una mayor presión antrópica a menores altitudes y por el aumento de mutaciones en altitudes más elevadas (mayor radiación UV), provocando una mayor diferenciación entre ambas zonas (Vera *et al.*, 2020). Estos aspectos se podrían analizar en estudios futuros, comparando poblaciones costeras con otras continentales en altitud elevada en distintas regiones ibéricas (Ramos-Capón *et al.*, 2024).

4.4.2 Análisis de diferenciación y estructura genética

El coeficiente de diferenciación interpoblacional (F_{ST}) obtenido a nivel global en este estudio fue de 0,044 lo que refleja un nivel bajo de diferenciación genética entre las localidades analizadas, aunque superior al valor obtenido con microsatélites (0,024), explicable por varianza interloci dado el bajo número disponible. El análisis con miles de marcadores SNPs a escala genómica ha permitido refinar el componente de diferenciación entre poblaciones, indicando un nivel cercano al límite entre divergencia baja y moderada (F_{ST} :0,05; Vilas & Álvarez, 2023).

Los valores de F_{ST} por pares (Tabla 13) indicaron en que las muestras poblacionales menos diferenciadas fueron Cíes-Barra ($F_{ST}= 0,014$) y las más diferenciadas fueron Barra- Sálvora ($F_{ST}=0,059$). Todos los valores de F_{ST} por pares fueron significativos ($p < 0,001$) señalando diferencias genéticas entre todas las poblaciones analizadas, pero con diferente magnitud relativa entre distintas comparaciones. La muestra C1 presentó la menor diferenciación con I1, con gradiente creciente respecto a las islas más al norte (I2 e I3). Por el contrario, Sálvora mostró la mayor diferenciación contra todas las restantes localidades, con el menor valor respecto a Ons. Los valores de F_{ST} más elevados, superiores a 0,5 (C1-I3, I1-I2 e I2-I3) indicarían una divergencia genética moderada entre esas localidades, sugiriendo un menor flujo génico entre dichas poblaciones (Vilas & Álvarez., 2023).

Tabla 13. Valores del coeficiente de diferenciación interpoblacional (F_{ST}) por pares de muestras poblacionales estudiadas. En la mitad superior se indica la significación: *** p-valor < 0,001.

Localidades	C1	I1	I2	I3
C1	-	***	***	***
I1	0,014	-	***	***
I2	0,043	0,056	-	***
I3	0,059	0,056	0,032	-

*** p-valor < 0,001

En base a estos resultados con SNPs la muestra continental en Barra (C1) mostró mayor diferenciación con las muestras insulares más alejadas, Ons y Sálvora. Mientras C1 resultó menos diferenciado de la muestra insular más cercana en Cíes, con una historia de

conectividad en la última glaciación (Arce Chamorro *et al.*, 2021). Estos resultados son concordantes con los detectados con marcadores microsatélites, a pesar de su número reducido, con particular interés del microsatélite Dro-42 como marcador de diferenciación interpoblacional en el área de estudio, de fácil implementación por PCR en muestras futuras.

Los valores de F_{ST} en este estudio se encuentran dentro del rango descrito con marcadores SNPs entre poblaciones de especies relacionadas de la misma familia. En particular, son similares a los encontrados entre pares de subpoblaciones con baja-moderada diferenciación de *Daphnimorpha capitellata* ($F_{ST}= 0,011-0,084$). Por otra parte, contrasta con los valores obtenidos en subpoblaciones con mayor grado de diferenciación *Daphnimorpha kudoii* ($F_{ST}= 0,082-0,152$) (Takahashi & Suyama, 2023).

En el trabajo realizado por Gajdošová *et al.*, (2024) en la especie *D. arbuscula*, describieron unos valores de F_{ST} por pares entre 0,0035 y 0,0079, indicando una divergencia muy baja y no significativa a escala microgeográfica, aunque con cierta subestructuración entre las poblaciones. Los valores de diferenciación interpoblacional pareados fueron superiores en *D. gnidium* en el presente estudio, congruente con la representación de muestras en distintos territorios insulares (Tabla 13).

Los análisis de agrupación bayesiana con STRUCTURE sugirieron que el número de unidades de población genéticamente homogéneas (K) más probable osciló entre dos y tres ($K= 2$; $K= 3$). Para $K= 2$ un primer grupo estaba compuesto por C1 (Barra) e I1 (Cíes) y el segundo grupo por I2 (Ons) e I3 (Sálvora). En $K= 3$ se refuerza el patrón de diferenciación mostrada para $K= 2$, agrupando C1 (Barra) con I1 (Cíes); pero detectando una ligera diferenciación entre I2 (Ons) y I3 (Sálvora). Finalmente, se analizó también $K= 7$ como valor sugerente de estabilización de probabilidad, donde se observó un patrón de difícil interpretación a escala local, pero que confirma la similitud entre C1 (Barra) e I1 (Cíes) en comparación con el resto de las áreas insulares, con ligera subestructuración entre Ons y Sálvora, coincidiendo con el análisis para $K= 3$.

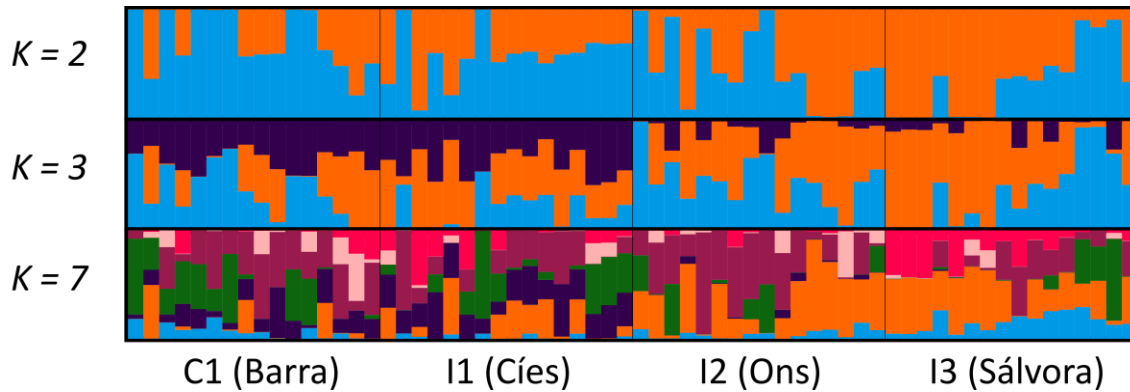


Figura 8. Agrupación genética representada por el análisis de STRUCTURE para K=2, K=3 y K= 7. Cada individuo está representado por una barra vertical, y la proporción cromática para cada barra se corresponde con la probabilidad posterior de asignación cada individuo a los diferentes grupos (K) estimado por el programa. Los individuos se agrupan por localidades de procedencia.

Los resultados seleccionados (Fig. 8) corresponden a los más probables detectados en el análisis STRUCTURE con distintos métodos. El gráfico de Probabilidades donde la curva se estabiliza sustentaría Ln (PD) para K= 7. Sin embargo, calculando el estadístico ΔK propuesto por Evanno *et al.* (2005), el valor más probable de K sería 2, seguido de 3, confirmando los resultados obtenidos (Fig. 8), en todo caso, revelando un patrón de estructuración genética concordante entre los distintos valores de K más probables. El valor de K sugerido por el análisis BIC resultó equivalente al valor óptimo del método de Evanno *et al.* (2005) mencionado.

Estos análisis de estructura respaldan los valores de diferenciación F_{ST} por pares (Tabla 14) indicando baja divergencia genética entre las localidades de Barra y Cíes, diferenciadas de Ons y Sálvora, también divergentes entre sí. Esto sugiere diferentes niveles de flujo génico y estructuración genética entre las poblaciones estudiadas.

La presencia de tres grupos diferenciados obtenidos en los análisis de STRUCTURE se confirmó con el diagrama DAPC. Los 13 componentes principales (PCs) permitieron abarcar el 53,8% de la varianza utilizando como análisis discriminatorio del *eigenvalue* 1 frente al 2 (Fig. 9). Los resultados obtenidos con SNPs en *D. arbuscula* descritos por Gajdošová *et al.*, (2024), revelaron un grado de subestructuración entre las diferentes

muestras poblacionales analizadas en una escala local microgeográfica, aunque la comparación entre distintas especies ha de ser tomada con precaución, bajo procesos evolutivos específicos.

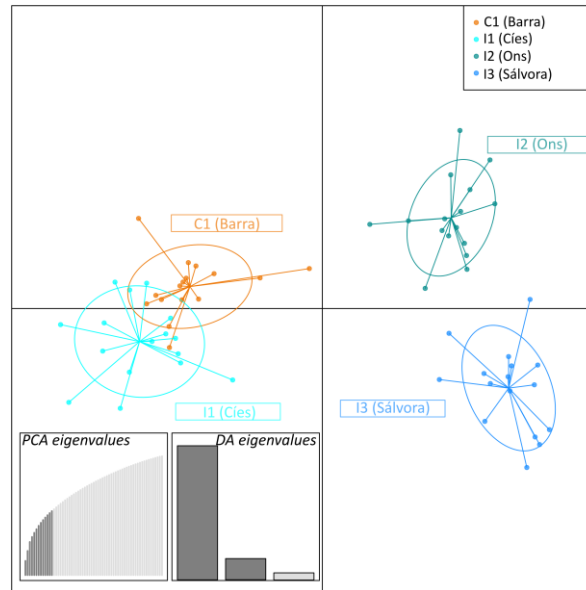


Figura 9. Análisis discriminante de los componentes principales (DAPC), representando el primer y segundo factor discriminante (DF) incluyendo el peso del análisis discriminante retenido (DA) para representar el 53,8% de la varianza

En resumen, los análisis de microsatélites mostraron una visión preliminar complementada con un amplio número de SNPs para obtener una imagen representativa a escala genómica de la diversidad y estructura genética de las poblaciones de *D. gnidium* en las costas atlánticas de Galicia. Los microsatélites evidenciaron una distribución heterogénea de la diversidad, con mayores valores en Barra (C1) y Cíes (I1), en comparación con las poblaciones aisladas de Ons (I2) y Sálvora (I3), donde se observa menor riqueza alélica y heterocigosis, sugiriendo efectos del aislamiento y la deriva. El análisis de estructuración con SNPs reveló una mayor diferenciación global entre localidades, con un 4,4% de la variación atribuible a diferencias entre poblaciones, indicando una separación entre las insulares y continentales.

La combinación de microsatélites y SNPs ha permitido identificar grupos poblacionales genéticamente diferenciados indicativos de diferentes unidades de gestión, de utilidad para apoyar planes de conservación para preservar la diversidad genética, en el marco del proyecto LIFE INSULAR (<https://www.lifeinsular.eu/>). Entre las orientaciones genética de gestión y conservación, tal como se ha descrito en diferentes plantas, se recomienda preservar la diversidad genética total detectada en el área de estudio, manteniendo el componente genético de diferenciación interpoblacional, para garantizar el potencial evolutivo de las poblaciones, evitando translocar material vegetal entre grupos genéticamente diferenciados (Vera *et al.* 2020; Bouza *et al.* 2023).

5. CONCLUSIONES

- La puesta a punto de un panel de nueve marcadores microsatélites en *Daphne gnidium* previamente caracterizados en *D. rodriguezii* demostró alta tasa de conservación en regiones cebadoras de ambas especies, aunque bajo rendimiento en calidad de genotipado y polimorfismo, desaconsejando su uso en estudios de variación genética en *D. gnidium*.
- Dos marcadores microsatélites (Dro-25 y Dro-42) resultaron viables e informativos en las muestras poblacionales objeto de análisis, confirmando su utilidad para estudios futuros en *D. gnidium*.
- La caracterización de marcadores SNP obtenidos mediante genotipado por secuenciación permitió disponer de un panel de 2011 loci de calidad con baja proporción de datos ausentes en las muestras poblacionales analizadas.
- El análisis preliminar de diversidad y diferenciación genética con marcadores microsatélites resultó altamente concordante con el análisis ampliado con SNPs. Los mayores niveles de diversidad genética se detectaron en el área insular de Cíes y continental de Barra, frente a niveles menores de diversidad en las islas de Ons y Sálvora, asociados con menores censos poblacionales y flujo génico reducido respecto al resto de poblaciones.

- Los análisis de diferenciación y estructura genética permitieron discriminar dos grupos insulares diferenciados en el Parque Nacional de Illas Atlánticas, la muestra poblacional de Cíes aparece más conectada al área continental en Barra que al resto de islas analizadas, Ons y Sálvora, a su vez relacionadas entre sí.
- La información genética proporcionada puede ser utilizada por las autoridades competentes para el desarrollo de acciones adecuadas de gestión y conservación, con el objetivo de preservar el potencial evolutivo de la especie y evitar la translocación de material vegetal entre grupos genéticamente diferenciados.

6. BIBLIOGRAFÍA

1. Arce Chamorro, C., Vidal Romaní, J.R. & Sanjurjo Sánchez, J. (2021). *Islas Cíes: una trampa eólica en la Ría de Vigo (Galicia, España) al final del último glaciar*. *GEOGACETA: Revista semestral de ciencias de la tierra / biannual journal of earth sciences*. 70, 3-6.
2. Banerjee, A., Dwivedi, M. D., Kasana, S., Singh, P., Kumar, V., Maity, D., & Pandey, A. K. (2022). *Phylogenetic relationships in Indian Daphne (Thymelaeaceae) based on nuclear ITS and cpDNA data*. *Biología*. 77, 3071-3086.
3. Besnier, F., & Glover, K. A. (2013). *ParallelStructure: A R Package to Distribute Parallel Runs of the Population Genetics Program STRUCTURE on Multi-Core Computers*. *PLoS ONE*. 8.
4. Bouza, C., Lorenzo, I., Rodríguez-Guitián, M. A., Casanova, A., Real, C., Romero, R., & Vera, M. (2023). *Genetic survey extension of the threatened Iberian Arnica montana L. revealed the presence of divergent plastid lineages and highly structured populations in northern Spain*. *Genetic Resources and Crop Evolution*. 70, 1677–1689.
5. Craig, J.M., & Bickmore, W.A. (1993). *Chromosome bands--flavours to savour*. *Bioessays*. 15, 349-354.
6. Crawford, A. M., Kappes, S. M., Paterson, K. A., deGotari, M. J., Dodds, K. G., et al. (1998). *Microsatellite evolution: testing the ascertainment bias hypothesis*. *J. Mol. Evol.* 46, 256–260.
7. Casanova, A., Maroso, F., Blanco, A., Hermida, M., Ríos, N., García, G., Manuzzi, A., Zane, L., Verissimo, A., García-Marín, J. L., Bouza, C., Vera, M., & Martínez, P. (2021). *Low impact of different SNP panels from two building-loci pipelines on RAD-Seq population genomic metrics: case study on five diverse aquatic species*. *BMC Genomics*. 22.
8. Chambers, E. A., Tarvin, R. D., Santos, J. C., Ron, S. R., Betancourth-Cundar, M., Hillis, D. M., Matz, M. v., & Cannatella, D. C. (2023). *2b or not 2b? 2bRAD is an effective alternative to ddRAD for phylogenomics*. *Ecology and Evolution*. 13.

9. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G. & Durbin, R. (2011). *The variant call format and VCFtools*. *Bioinformatics*. 1000 Genomes Project Analysis Group. 1; 27, 2156-2158.
10. Ellegren, H. (2000). *Microsatellite mutations in the germline: implications for evolutionary inference*. *Trends Genet.* 16, 551-558.
11. Ellegren, H., Primmer, C. R. & Sheldon, B. C. (1995). *Microsatellite 'evolution': Directionality or bias?* *Nat. Genet.* 11, 360–362.
12. El Mouzazi, I., Bour, A., & Khal-Layoun, S. (2023). *Phytochemical and biological activities of Daphne gnidium L. species: A systematic review*. In *Acta Ecologica Sinica*. Ecological Society of China. 44; 3, 428-443.
13. 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.
14. Falush, D., Stephens, M., & Pritchard, J. K. (2007). *Inference of population structure using multilocus genotype data: Dominant markers and null alleles*. *Molecular Ecology Notes*. 7, 574–578.
15. Gajdošová, Z., Šlenker, M., Svitok, M., Šrámková, G., Blanár, D., Cetlová, V., Kučera, J., Turisová, I., Turis, P., & Slovák, M. (2024). *Unravelling some factors affecting sexual reproduction in rock-specialist shrub: Insight from an endemic Daphne arbuscula (Thymelaeaceae)*. *PLoS ONE*. 19.
16. García-Verdugo, C., Sajeva, M., la Mantia, T., Harrouni, C., Msanda, F., & Caujapé-Castells, J. (2015). *Do island plant populations really have lower genetic variation than mainland populations? Effects of selection and distribution range on genetic diversity estimates*. *Molecular Ecology*. 24, 726–741.
17. García-Verdugo, C., Illera, J. C., & Traveset, A. (2019). *Characterization of microsatellite markers for the endangered Daphne rodriguezii (Thymelaeaceae) and related species*. *Applications in Plant Sciences*. 7, 1–5.
18. Goudet, J. (2003). *FSTAT (Version 2.9.3.2): A program to estimate and test gene diversities and fixation indices*. Updated from Goudet (1995).
19. Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). *dartr: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing*. *Molecular Ecology Resources*. 18, 691–699.
20. Halda-Alija, L., Hendricks, S. P., & Johnston, T. C. (2001). *Spatial and temporal variation of Enterobacter genotypes in sediments and the underlying hyporheic zone of an agricultural stream*. *Microbial Ecology*. 42, 286–294.
21. Jacob, H.J., Lindpaintner, K., Lincoln, S.E., Kusumi, K., Bunker, R.K., Mao, Y.P., Ganten, D., Dzau, V.J. & Lander, E.S. (1991). *Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat*. *Cell*. 4; 67, 213-224.
22. Jombart, T. (2008). *Adegenet: A R package for the multivariate analysis of genetic markers*. *Bioinformatics*. 24, 1403–1405.
23. Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). *DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors*. *Methods in Ecology and Evolution*. 4, 782–788.

24. Khouchlaa, A., el Menyiy, N., Guaouguaou, F. E., el Baaboua, A., Charfi, S., Lakhdar, F., el Omari, N., Taha, D., Shariati, M. A., Rebezov, M., El-Shazly, M., & Bouyahya, A. (2021). *Ethnomedicinal use, phytochemistry, pharmacology, and toxicology of Daphne gnidium: A review. Journal of Ethnopharmacology*. 275, 114-124.
25. 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.
26. 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.
27. Maroso, F., Vera, M., Ferreiro, J., Mayol, M., Riba, M., Ramil-Rego, P., Martínez, P., & Bouza, C. (2021). *Genetic diversity and structure of Taxus baccata from the Cantabrian-Atlantic area in northern Spain: A guide for conservation and management actions. Forest Ecology and Management*. 482.
28. Pembleton, L. W., Cogan, N. O. I., & Forster & J. W. (2013). *StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Molecular Ecology Resources*. 13, 946–952.
29. Pérez-González, J., Carranza, J., Anaya, G., Broggin, C., Vedel, G., de la Peña, E., & Membrillo, A. (2023). *Comparative Analysis of Microsatellite and SNP Markers for Genetic Management of Red Deer. Animals*. 13.
30. Pierce, B.A. (2015). *Genética: un enfoque conceptual*. 5ª ed. Madrid: Médica Panamericana.
31. P.O.W.O. (2024). *Solanum lycopersicum L. Plantas del mundo en línea*.
32. Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) *Inference of population structure using multilocus genotype data. Genetics*. 155, 945-959.
33. Ramos-Capón, C., Vargas, P., Pomedá-Gutiérrez, F., & Martín-Hernanz, S. (2024). *Time-calibrated phylogeny of Daphne (Thymelaeaceae): Pre-mediterranean temporal origin of the sclerophyllous Daphne gnidium. Journal of Systematics and Evolution*. 62, 20–37.
34. Roccotiello, E., Casazza, G., Galli, L., Cornara, L., Moncalvo, A., & Minuto, L. (2009). *The flower biology of Daphne gnidium L. (Thymelaeaceae). Plant Systematics and Evolution*. 279, 41–49.
35. Rousset, F., Leblois, R., & Cournet, J. (2008). *Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Molecular Ecology Resources*. 8, 103-106.
36. Rubio de Casas, R. (2006). *Microevolución de esclerófilas mediterráneas: diferenciación fenotípica y genética en Quercus coccifera L. y Olea europaea L. Tesis (Doctoral), E.T.S.I. Agrónomos (UPM)*. 289, 30-45.
37. Takahashi, D. & Suyama, Y. (2023). *Demographic history and conservation aspects of two relict shrubs (Daphnimorpha spp., Thymelaeaceae) showing disjunct and restricted distribution in the southern part of Japan. Tree Genetics & Genomes*. 19, 43.
38. Vera, M., Mora, G., Rodríguez-Gutián, M. A., Blanco, A., Casanova, A., Real, C., Romero, R., & Bouza, C. (2020). *Living at the edge: population differentiation in endangered Arnica montana from NW Iberian Peninsula. Plant Systematics and Evolution*. 306, 1–14.
39. Vilas, R., & Alvarez, G. (2022). *Genética evolutiva*. Editorial Síntesis.

40. Wang, S., Meyer, E., McKay, J.K. & Matz, M.V. (2012). *2b-RAD: a simple and flexible method for genome-wide genotyping*. *Nat Methods*. 20; 9, 808-10.
41. Weir BS & Cockerham CC. (1984). *Estimating f-statistics for the analysis of population structure*. *Evolution*. 38, 1358-1370.
42. Zane, L., Bargelloni, L. & Patarnello, T. (2002). *Strategies for microsatellite isolation: a review*. *Mol Ecol*. 11, 1-16.
43. Zimmerman, S. J., Aldridge, C. L., & Oyler-Mccance. (2020). *An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern*. *BMC Genomics*. 21.