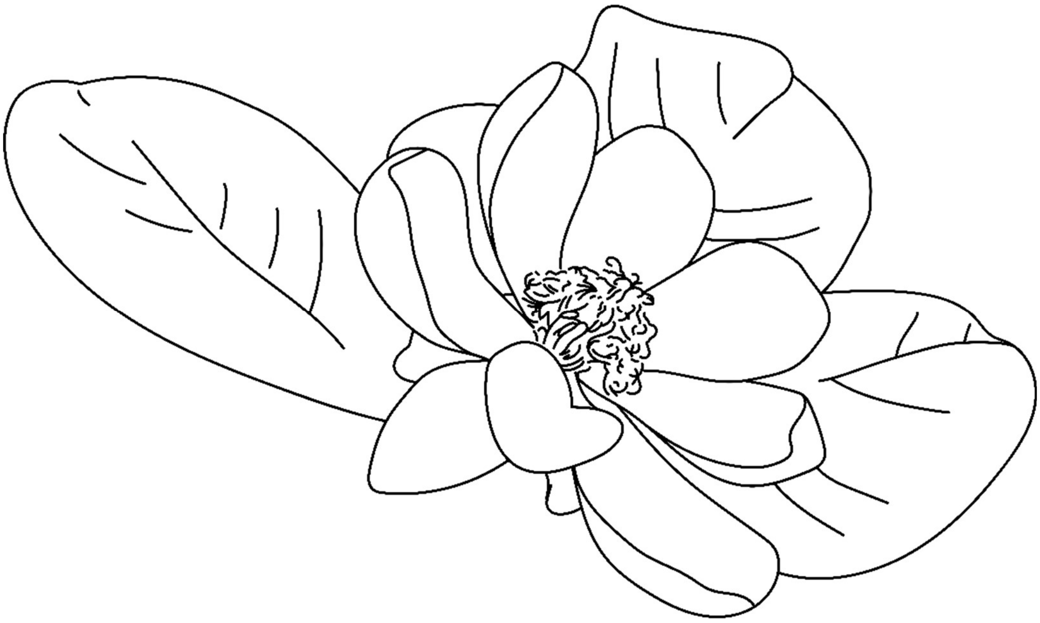
An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola



Tim Claerhout
Student number: 01601048

Supervisor: Prof. Dr. Lars Chatrou
Counsellors: Dr. Emily Veltjen, Dr. Diederik Strubbe

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Science in Biology.

Academic year: 2020 – 2021

Word count: 20.200

© All rights reserved. This thesis contains confidential information and confidential research results that are property to the UGent. The contents of this master thesis may under no circumstances be made public, nor complete or partial, without the explicit and preceding permission of the UGent representative, i.e. the supervisor. The thesis may under no circumstances be copied or duplicated in any form, unless permission granted in written form. Any violation of the confidential nature of this thesis may impose irreparable damage to the UGent. In case of a dispute that may arise within the context of this declaration, the Judicial Court of Gent only is competent to be notified.

© Deze masterproef bevat vertrouwelijke informatie en vertrouwelijke onderzoeksresultaten die toebehoren aan de UGent. De inhoud van de masterproef mag onder geen enkele manier publiek gemaakt worden, noch geheel noch gedeeltelijk zonder de uitdrukkelijke schriftelijke voorafgaandelijke toestemming van de UGent-vertegenwoordiger, in casu de promotor. Zo is het nemen van kopieën of het op eender welke wijze dupliceren van het eindwerk verboden, tenzij met schriftelijke toestemming. Het niet respecteren van de confidentiële aard van het eindwerk veroorzaakt onherstelbare schade aan de UGent. Ingeval een geschil zou ontstaan in het kader van deze verklaring, zijn de rechtbanken van het arrondissement Gent uitsluitend bevoegd daarvan kennis te nemen.

This thesis should be cited as: Claerhout, T. (2021). An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola (Unpublished master's thesis). Ghent University, Belgium. pp. 58.

Cover art: *Magnolia domingensis,* Robbe Strybol. Based on a photograph from Ramón Elías Castillo Torres.

Table of Contents

Objectives and Research Hypotheses

In this master's dissertation, we aim to deliver insights regarding the distribution and genetic health (i.e. the amalgamation between genetic diversity, gene flow and geographic extent of a species and/or population) of four endangered *Magnolia* species (Magnoliaceae) from the island of Hispaniola. We strive to formulate concrete suggestions for each species' future conservation efforts. Using molecular data, we investigate their genetic diversity by employing conservation genetic analyses. Furthermore, using species distribution modelling, a popular modelling framework in conservation biology, we predict the species' potential distribution and habitat suitability across Hispaniola. Seven null hypotheses (**H_X**) across three objectives were tested:

1) To predict suitable *Magnolia* habitat for conservation efforts and to guide explorations for new populations.
    - **H_1**: Suitable *Magnolia* habitat is only present in the mountain ranges in which they reside.
    - **H_2**: IUCN Red List Status is correlated with the amount of suitable *Magnolia* habitat in the vicinity.
2) To assess the population structure to infer conservation units.
    - **H_3**: Genetic structuring of the Magnolias of the Dominican Republic follows morphological species delineation as described by Howard (1948).
    - **H_4**: Each sampled locality represents one genetic cluster, as found for the majority of *Magnolia* species in Veltjen et al. (2019).
3) To quantify genetic diversity, enabling us to evaluate the genetic health of the populations.
    - **H_5**: Given the small population sizes, their endemic and (Critically) Endangered status, the *Magnolia* populations of the Dominican Republic have inbreeding coefficients that significantly differ from zero.
    - **H_6**: Genetic diversity is correlated with the IUCN Red List Status.
    - **H_7**: Genetic diversity is correlated with the amount of suitable *Magnolia* habitat in the vicinity.

Executive Summaries (EN/NL/ES)

Summary

On the island of Hispaniola, five endemic species of *Magnolia* occur, all of which are threatened with extinction: *Magnolia domingensis, M. ekmanii, M. emarginata, M. hamorii* and *M. pallescens.* The most prevalent threats include habitat fragmentation, illegal logging, land conversion for agriculture and livestock use, forest fires and little natural regeneration. Current knowledge regarding the distribution and genetic health (i.e. the amalgamation between genetic diversity, gene flow and geographic extent of a species and/or population) of the Magnolias of Hispaniola is extremely scarce, hampering targeted conservation actions. To define conservation units and their respective threats, information on the distribution of the focal species is essential. Species distribution modelling (SDM) is a popular modelling framework that allows us to predict the Magnolias' potential distribution and quantify its habitat suitability in Hispaniola. Furthermore, genetic information on the focal species enables conservation practitioners to answer key questions regarding their management. One preliminary study already reported high genetic structuring, yet little inbreeding for four of the five *Magnolia* species in Hispaniola. The objective of this study is to deliver data on the potential distribution and the genetic health of the Magnolias of Hispaniola to state concrete guidelines for effective conservation management. More specifically, we predicted suitable *Magnolia* habitat for conservation efforts and to guide explorations for new populations using SDM. Furthermore, applying conservation genetics, we assessed the genetic structure and quantified genetic diversity parameters. This allowed us to make statements about the genetic health of the populations. By integrating these results, we inferred conservation units and proposed conservation recommendations that strive to ensure a sustainable future for the Magnolias of Hispaniola. Firstly, using SDM, we analysed 21 variables describing climate and landscape features and fitted eight modelling algorithms in an ensemble framework to predict the potential distribution for the Magnolias of Hispaniola. These predictions were summarized in three final distribution maps. The distribution of potential habitats was restricted to the various mountain ranges on Hispaniola. Secondly, 16 microsatellite markers were employed to test the genetic structure and degree of inbreeding for 417 individuals across three of the five species: *M. domingensis*, *M. hamorii* and *M. pallescens*. Diversity statistics were calculated for each of those three species' populations and localities. Analyses for genetic structure showed strong species integrity and moderate genetic differentiation between population pairs of *M. domingensis* and *M. pallescens*. Consequently, five populations across three species were defined. For *M. pallescens*, significant inbreeding was retrieved in the population Ébano Verde; and the localities "Entrance" and "Montellano." The population Loma Rodríguez (*M. domingensis*) had the lowest genetic diversity. Combining all results, we designated each species as one conservation unit. *Magnolia domingensis* is highlighted as it contains the lowest genetic diversity, the least number of (known) individuals and the lowest number of sampling localities. We conclude that the three genetically investigated Magnolias consist of five populations in total, with pronounced genetic structuring found only in *M. pallescens*. Ample genetic diversity is present with minimal inbreeding. Much potential *Magnolia* habitat in Hispaniola is unexplored, which might contain yet undiscovered *Magnolia* individuals. For conservation management, we propose concrete species-specific actions in terms of exploration, protection and reinforcement. The main recommendation includes protection and/or instatement of habitat corridors between populations of *M. domingensis* and *M. ekmanii*, whereby the SDM results will serve as a guideline for spatial prioritization.

Samenvatting

Op het eiland Hispaniola komen vijf endemische soorten *Magnolia* voor, die allen met uitsterven bedreigd zijn: *Magnolia domingensis*, *M. ekmanii, M. emarginata, M. hamorii* and *M. pallescens.* De belangrijkste bedreigingen zijn habitat fragmentatie, illegale houtkap, land conversie in functie van landbouw en veeteelt, bosbrand en een tekort aan natuurlijke verjonging. De huidige kennis omtrent de verspreiding en genetische toestand van de Magnolias van Hispaniola is zeer beperkt, wat gerichte conservatie praktijken belemmert. Om conservatie eenheden en hun respectievelijke bedreigingen te bepalen is informatie over de verspreiding van de doelsoorten essentieel. “Species distribution modelling” (SDM) is een populaire modelleertechniek die ons toelaat om de potentiële verspreiding en de geschiktheid van het habitat van de *Magnolia* soorten in Hispaniola te voorspellen. Bovendien laat de genetische informatie toe om essentiële vragen omtrent het beheer van de doelsoorten te beantwoorden. Eén preliminaire studie rapporteerde reeds een uitgesproken genetische structuur, maar weinig inteelt voor vier van de vijf *Magnolia* soorten in Hispaniola. Deze studie doelt op het leveren van data over de potentiële verspreiding en de genetische toestand van de Magnolias van Hispaniola. Zo kunnen we concrete richtlijnen opstellen voor toekomstige, doelgerichte beheersmaatregelen. Specifiek hebben we geschikt *Magnolia* habitat voorspeld voor toekomstige conservatie praktijken en expedities die zullen zoeken naar nieuwe populaties. Bovendien, met behulp van conservatie genetica, hebben we de genetische structuur beoordeeld en parameters van genetische diversiteit berekend. Dit liet ons toe om uitspraken te doen over de genetische toestand van de populaties. Door deze resultaten te integreren, hebben we conservatie eenheden bepaald en aanbevelingen gedaan inzake conservatie beheer dat streeft naar een duurzame toekomst voor de Magnolias van Hispaniola. Ten eerste, met behulp van SDM analyseerden we 21 variabelen met betrekking tot klimaat en landschapseigenschappen om er vervolgens acht modellen op toe te passen en zo de potentiële verspreiding van de Magnolias van Hispaniola te voorspellen. Deze voorspellingen werden samengevat in drie finale verspreidingskaarten. De verspreiding van het geschikte habitat was beperkt tot de diverse bergketens op Hispaniola. Ten tweede, met behulp van 16 microsatelliet merkers, testten we de genetische structuur en de mate van inteelt voor 417 individuen over drie van de vijf soorten: *M. domingensis*, *M. hamorii* and *M. pallescens*. Diversiteitsindices werden berekend voor elke populatie en locatie van de drie soorten. Analyses met betrekking tot de genetische structuur toonden een sterke soortintegriteit en middelmatige genetische differentiatie tussen populatie paren van *M. domingensis* en *M. pallescens.* Daarbij werden vijf populaties verspreid over drie soorten vastgesteld. Voor *M. pallescens* werd significante inteelt gevonden voor de populatie Ébano Verde en de locaties “Entrance” en “Montellano.” De populatie Loma Rodríguez (*M. domingensis*) vertoonde de laagste genetische diversiteit. Door alle resultaten te integreren werd elke soort als één conservatie eenheid aangewezen. *Magnolia domingensis* werd aangewezen als het meest kwetsbaar aangezien het de laagste genetische diversiteit vertoont en het laagste aantal (gekende) individuen en locaties heeft. We concluderen dat de drie *Magnolia* soorten die genetisch bestudeerd zijn in totaal uit vijf populaties bestaan, waarvan enkel *M. pallescens* een uitgesproken genetische structuur vertoont. Verder is nog genoeg genetische diversiteit aanwezig en is inteelt slechts minimaal aanwezig. Veel potentieel geschikt *Magnolia* habitat in Hispaniola is nog steeds niet geëxploreerd. Hier zouden nog onontdekte *Magnolia* individuen aanwezig kunnen zijn. Voor conservatie beheer stellen we soortspecifieke maatregelen voor op het gebied van exploratie, bescherming en versterking. De voornaamste aanbevelingen omvatten het beschermen en/of oprichten van habitat corridors tussen populaties van *M. domingensis* en *M. ekmanii*, waarbij de SDM-resultaten kunnen dienen als een richtlijn voor het stellen van ruimtelijke prioriteiten.

Resumen

Se encuentran cinco especies endémicas de *Magnolia* en la isla La Española, todas en peligro de extinción: *Magnolia domingensis*, *M. ekmanii*, *M. emarginata*, *M. hamorii* y *M. pallescens*. Las amenazas más frecuentes incluyen fragmentación del hábitat, tala ilegal, conversión de tierras para uso agrícola y ganadero, incendios forestales y poca regeneración natural. El conocimiento actual sobre la distribución y la salud genética de las Magnolias de Hispaniola es extremadamente escaso, lo cual dificulta las acciones de conservación específicas. Para definir las unidades de conservación y sus respectivas amenazas, la información sobre la distribución de las especies focales es fundamental. El modelado de distribución de especies (MDE) es un marco de modelado popular que nos permite predecir la distribución potencial de las Magnolias y cuantificar la idoneidad de su hábitat en La Española. Además, la información genética sobre las especies focales permite a los profesionales de la conservación responder preguntas clave sobre su gestión. Un estudio preliminar ya informó una alta estructuración genética, pero poca endogamia para cuatro de las cinco especies de *Magnolia* en La Española. El objetivo de este estudio es presentar datos sobre la distribución potencial y la salud genética de las Magnolias de La Española para establecer pautas concretas para una gestión de conservación eficaz. Más específicamente, predijimos un hábitat de *Magnolia* adecuado para los esfuerzos de conservación y para guiar las exploraciones de nuevas poblaciones utilizando MDE. Además, aplicando la genética de la conservación, evaluamos la estructura genética y cuantificamos los parámetros de diversidad genética. Esto nos permitió presentar información sobre la salud genética de las poblaciones. Al integrar estos resultados, inferimos unidades de conservación y propusimos recomendaciones de conservación con el fin de asegurar un futuro sostenible para las Magnolias de La Española. En primer lugar, utilizando MDE, analizamos 21 variables que describen las características del clima y el paisaje y ajustamos ocho algoritmos de modelado en un marco de conjunto para predecir la distribución potencial de las Magnolias de La Española. Estas predicciones se resumieron en tres mapas de distribución finales. La distribución de hábitats potenciales se restringió a las diversas cadenas montañosas de La Española. En segundo lugar, se emplearon 16 marcadores de microsatélites para probar la estructura genética y el grado de endogamia de 417 individuos en tres de las cinco especies: *M. domingensis*, *M. hamorii* y *M. pallescens*. Se calcularon estadísticas de diversidad para cada una de las poblaciones y localidades de esas tres especies. Los análisis de la estructura genética mostraron una fuerte integridad de las especies y una diferenciación genética moderada entre los pares de poblaciones de *M. domingensis* y *M. pallescens*. En consecuencia, se definieron cinco poblaciones de tres especies. Para *M. pallescens*, se obtuvo una endogamia significativa en la población Ébano Verde; y las localidades “Entrance” y “Montellano”. La población Loma Rodríguez (*M. domingensis*) tuvo la menor diversidad genética. Combinando todos los resultados, designamos a cada especie como una unidad de conservación. *Magnolia domingensis* se destaca porque contiene la menor diversidad genética, el menor número de individuos (conocidos) y el menor número de localidades de muestreo. Concluimos que las tres Magnolias investigadas genéticamente consisten en cinco poblaciones en total, con una estructura genética pronunciada encontrada solo en *M. pallescens*. Existe una amplia diversidad genética con una mínima consanguinidad. Gran parte del hábitat potencial de *Magnolia* en La Española está inexplorado, lo cual podría contener individuos de *Magnolia* aún no descubiertos. Para la gestión de la conservación, proponemos acciones concretas para especies específicas en términos de exploración, protección y refuerzo. La recomendación principal incluye la protección y/o establecimiento de corredores de hábitat entre las poblaciones de *M. domingensis* y *M. ekmanii*, por lo que los resultados del MDE servirán como guía para la priorización espacial.

Master's thesis – Tim Claerhout – 2021
Master of Science in Biology – Major Biodiversity & Evolution

An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola

Tim Claerhout · Emily Veltjen · Diederik Strubbe · Ramón Elías Castillo Torres · Marie-Stéphanie Samain · Lars Chatrou

Received: 01/06/2021

Abstract On the island of Hispaniola, five endemic species of *Magnolia* occur, all of which are threatened with extinction. Little is known about their distribution and genetic health, hampering targeted conservation actions. One preliminary study already reported high genetic structuring, but little inbreeding. The objective of this study is to deliver data on the potential distribution and the genetic health of the Magnolias of Hispaniola to state concrete guidelines for effective conservation management. Using species distribution modelling (SDM), we analyzed 21 variables describing climate and landscape features to predict the potential distribution of the Magnolias of Hispaniola. Furthermore, 16 microsatellite markers were employed to test the genetic structure and degree of inbreeding for 417 individuals across three of the five species. The distribution of potential habitats was restricted to the various mountain ranges on Hispaniola. Analyses for genetic structure showed strong species integrity and moderate genetic differentiation between population pairs among two species. We conclude that the Magnolias of Hispaniola consist of five populations with ample genetic diversity and only minimal inbreeding. Moreover, much potential habitat is unexplored, which might contain new *Magnolia* individuals. For conservation management, we propose concrete species-specific actions in terms of exploration, protection, and reinforcement. The main recommendation includes protection and/or instatement of habitat corridors between populations of *M. domingensis* and *M. ekmanii*, whereby these SDM results will serve as a guideline for spatial prioritization.

Resumen Se encuentran cinco Magnolias endémicas en la isla La Española, , todas en peligro de extinción. Se sabe poco sobre su distribución y salud genética, dificultando las acciones de conservación específicas. Un estudio preliminar reportó una estructura genética alta y poca endogamia. El objetivo de este estudio es presentar datos sobre distribución potencial y salud genética de las Magnolias de La Española para establecer pautas concretas para gestión de conservación eficaz. Utilizando modelos de distribución de especies (MDE), analizamos 21 variables que describen las características del clima y el paisaje para predecir su distribución potencial. Además, se emplearon 16 marcadores de microsatélites para pro-

---

Tim Claerhout
Department of Biology, Systematics and Evolutionary Botany lab, Ghent University, K.L. Ledeganckstraat 35, Ghent, B-9000, België
E-mail: Tim.Claerhout@UGent.be

Emily Veltjen
Department of Biology, Systematics and Evolutionary Botany lab, Ghent University, K.L. Ledeganckstraat 35, Ghent, B-9000, België

Diederik Strubbe
Department of Biology, Terrestrial Ecology Unit, Ghent University, K.L. Ledeganckstraat 35, Ghent, B-9000, België

Ramón Elías Castillo Torres
Fundación PROGRESSIO y Jardin Botánico Nacional Dr. Rafael M. Moscoso, Dominican Republic

Marie-Stéphanie Samain
Red de Diversidad Biológica del Occidente Mexicano, Insti- tuto de Ecología, Avenida Lázaro Cárdenas 253, Pátzcuaro, 61600, Michoacán, Mexico
Department of Biology, Systematics and Evolutionary Botany lab, Ghent University, K.L. Ledeganckstraat 35, Ghent, B-9000, België

Lars Chatrou
Department of Biology, Systematics and Evolutionary Botany lab, Ghent University, K.L. Ledeganckstraat 35, Ghent, B-9000, België
E-mail: Lars.Chatrou@UGent.be

bar la estructura genética y el grado de endogamia de 417 individuos en tres de las cinco especies. La distribución de hábitats potenciales se restringió a las cadenas montañosas de La Española. Los análisis de estructura genética mostraron una fuerte integridad de las especies y una diferenciación genética moderada entre pares de poblaciones entre dos especies. Concluimos que las tres Magnolias genéticamente investigadas consisten en cinco poblaciones en total, con una amplia diversidad genética y una mínima consanguinidad. Además, gran parte del hábitat potencial está inexplorado, pudiendo contener individuos de *Magnolia* aún no descubiertos. Para la gestión de la conservación, proponemos acciones concretas para especies específicas en términos de exploración, protección y refuerzo. La recomendación principal incluye protección y/o establecimiento de corredores de hábitat entre las poblaciones de *M. domingensis* y *M. ekmanii*, por lo que estos resultados del MDE servirán como guía para la priorización espacial.

Fig. 1. Tropical montane cloud forest, the typical habitat of *Magnolia* in the Caribbean. Photograph: Emily Veltjen.

Keywords Species distribution modelling · conservation genetics · Magnoliaceae · Hispaniola · microsatellites · genetic diversity

Introduction

Tropical montane cloud forests (henceforth TMCFs) are a rare, but extremely biodiverse habitat. Although TMCFs only comprises 0.26% of the Earth's surface, they are home to a disproportionate amount of species (Bubb et al. 2004). This habitat is characterized by heavy rainfall, high species endemism and low resilience to disturbance (Hamilton et al. 1995). Despite its importance as a vital source of water and its high concentrations of biodiversity, TMCFs face similar threats as other tropical habitats: habitat fragmentation, climate change, illegal logging, etc. (Bubb et al. 2004). Since the 1993 international TMCF symposium (Hamilton et al. 1995), many challenges still need to be overcome to guarantee sustainable conservation of TMCFs.

One island that comprises this rare habitat is Hispaniola (Fig. 1), the second largest island in the Caribbean sea. It is divided in two nations, Haiti and the Dominican Republic (henceforth DR). While the former is infamous for being one of the most deforested countries on earth, with less than 1% of its original primary forests remaining (Hedges et al. 2018), the latter likewise suffers from forest loss and degradation (Sangermano et al. 2015a). Nonetheless, Hispaniola remains a biodiversity hotspot due to its high species endemism, housing about 6000 endemic plant species (Maunder et al. 2008; Fajardo et al. 2016; Cano-Ortiz et al. 2016). Five of these are endemic and threatened *Magnolia* species, *M. emarginata*, *M. ekmanii*, *M. domingensis*, *M. pallescens* and *M. hamorii* (Castillo 2016), consisting of only a few populations each (Fig. 2). Consequently, the Magnolias of Hispaniola are in great risk of extinction with one of the most prevalent factors being habitat loss due to timber exploitation, land conversion and other destructive activities (Castillo et al. 2018; Veltjen et al. 2019). As these species reside in TMCFs, their survival is intricately linked to the preservation of this habitat. Moreover, *Magnolia* is an eye-catching genus due to their beautiful flowers, aromatic leaves and ornamental use, which gives it great potential as an umbrella species for conservation (Roberge and Angelstam 2004). Hence, conservation of these species warrants the conservation of their habitat.

To employ effective conservation strategies, several sources of information are essential: threats, taxonomy, conservation units, the species biology, geographical extent, and the underlying genetic diversity (Kramer and Havens 2009; IUCN 2017). Current knowledge regarding the distribution and genetic diversity of the Magnolias of Hispaniola is extremely scarce (Cires et al. 2013; Castillo et al. 2018). We aim to resolve this issue by integrating two research fields, species distribution modelling (SDM) and conservation genetics, in forthcoming conservation strategies. This integration allows us to identify and set priorities for management of these species, to improve future decision making and to ensure a sustainable future as such.

A first problem arises when considering the geographic extent of the species. This information is only scarcely available for the Magnolias of Hispaniola (Castillo et al. 2018; Veltjen et al. 2019), which is exemplar of the Wallacean shortfall (i.e. the incomplete knowledge of species distributions) in tropical countries (Urbina-Cardona et al. 2019). However, to define conservation units and their threats, this information is essential. Species distribution modelling is a popular modelling framework which is increasingly used to address questions in conservation biology, ecology and evolution (Guisan and Thuiller 2005; Guillera-Arroita et al. 2015). It allows us to predict the (potential) species distribution and the habitat suitability in Hispaniola, based on species occurrence data and environmental information. These insights set the scene for future explorations to potentially find new populations and provides a guideline for spatial prioritisation of conservation actions.

Other than a species' geographical extent, its genetic information enables us to solve key questions regarding management of threatened species. Conservation genetics aims to reduce the risk of extinction in endangered species by quantifying and investigating their genetic diversity (Frankham et al. 2015). It strives to safeguard species as dynamic entities with sufficient evolutionary potential over the long term. A preliminary study of Veltjen et al. (2019) assessed the population structure and genetic health (i.e. the amalgamation between genetic diversity, gene flow and geographic extent of a species) for all Caribbean *Magnolia* species, including four out of the five species from Hispaniola. *Magnolia emarginata* was not assessed as no individuals of this Haitian species were found. The authors reported little inbreeding, yet high genetic structuring. As this study included only twenty individuals per population, which is only a fraction of the known number of populations and individuals of the *Magnolia* diversity in Hispaniola, we aim to study the genetic diversity of *M. pallescens*, *M. domingensis* and *M. hamorii* in depth. With this information, we can propose concrete conservation actions in terms of exploration, protection and reinforcement to minimize future inbreeding with the goal of maximizing the species' genetic diversity over the long term. By doing so, we strive to maintain healthy populations with enough evolvability to make them flexible and/or resistant to future disturbances.

Here, we aim to resolve urgent questions regarding the distribution and genetic health of the Magnolias of Hispaniola with the intention to make concrete suggestions for each species' future conservation efforts. We specifically wish to (1) predict suitable *Magnolia* habitat for conservation efforts and to guide explorations for new populations using SDM. Furthermore, using conservation genetics, we will (2) assess the population structure to infer conservation units and (3) quantify genetic diversity, enabling us to make statements about the genetic health of the populations. As a result, the proposed conservation actions strive to save these species from extinction.

Materials and Methods

Species Distribution Modelling

SDMs are capable of generating predictions of species' potential distributions. It consists of two main steps: data preparation and model fitting.

Data preparation allows the data to be preprocessed. First, occurrence data need to be rarefied over a predefined rarefying distance to minimize spatial autocorrelation. Available occurrence data are often not gathered using systematic surveying methods and do not represent a random sampling across a species' range (Phillips et al. 2009). Occurrence data are thus often biased in terms of sampling effort. Without rarefaction, predictions will be overestimated in environments that are similar to the more intensely sampled areas. Second, species absences somehow need to be incorporated in SDM algorithms, which are typically only rarely available (Zurell 2020). To tackle this lack of reliable absence data, "pseudo-absences" (also referred to as background data) are generated from a predefined background area (Barbet-Massin et al. 2012). The number of pseudoabsences to choose, relative to the number of presences (i.e. the prevalence ratio), has led to much discussion, but mainly depends on the SDM algorithm (Liu et al. 2019). Therefore, we tested different pseudo-absence selection strategies in accordance with Lobo et al. (2010).

Model fitting follows after data preparation. As algorithms have different underlying assumptions and extrapolate to the environment in a different manner, it has been proposed to combine predictions across multiple modelling methods. This is called ensemble modelling and enhances predictive performances (e.g. Hao et al. (2020)). We built and evaluated the performance of the modelling algorithms by cross-validating the data. This model fitting and evaluation strategy divides the data in two subsets: a training set, which trains the respective model, and a test set, which tests its predictive accuracy. Finally, an ensemble of SDM models that answer to a predefined minimum predictive accuracy was used to obtain final pre-

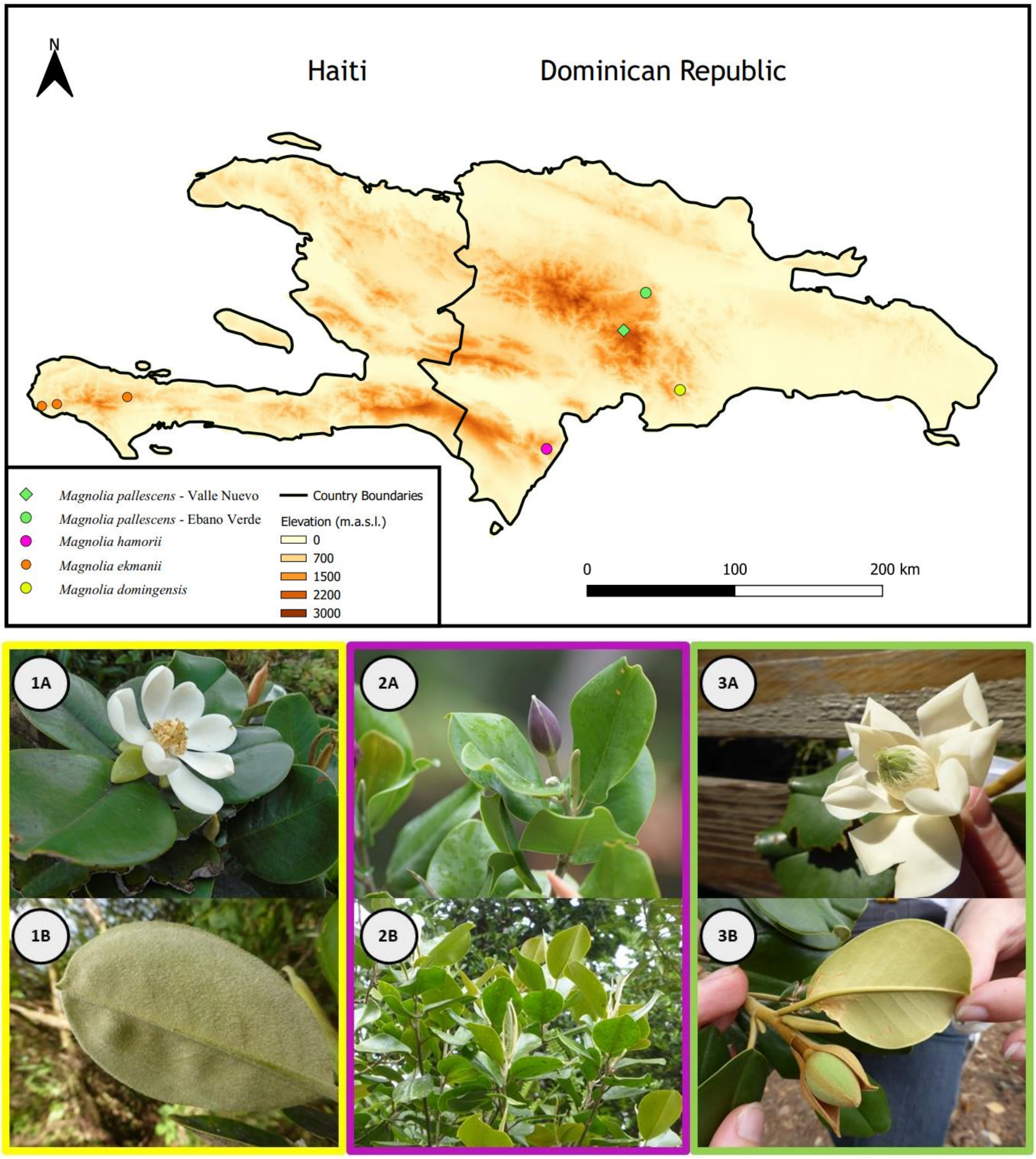


Fig. 2. Map visualizing the five *Magnolia* species from Hispaniola with additional morphological illustration. Base layer depicts elevation as meters above sea level (m.a.s.l.). Classification follows Figlar and Nooteboom (2004). **1**. *Magnolia domingensis*. A. Flower in male phase. B. Villose leaves. **2**. *Magnolia hamorii*. A. Closed flower bud. B. Leaves and evidence of conduplicate leaf prefoliation. **3**. *Magnolia pallescens*. A. Immature flower with underdeveloped reproductive structures. B. Elliptic leaf with abaxial sericeous pubescence. Photographs: 1A & 2A – Ramon Elías Castillo Torres; 1B, 2B, 3A, 3B – Emily Veltjen.

dictions of a species habitat suitability and potential geographic distribution across the study area.

Despite our focus on Magnolias from the Dominican Republic, SDM was executed on Hispaniola as a whole (i.e. including Haiti) as delineating study areas based on ecologically meaningless criteria such as country borders is likely to result in spurious predictions of habitat suitability (Bystriakova et al. 2012).

Occurrence Data and Predictor Variables

Occurrence data of four of the five species from Hispaniola were obtained from Veltjen (2020), Castillo et al. (2018), herbarium material (Online Resource 1) and pers. comm. of Joel Timyan (Société Audubon Haiti). Occurrence data were retained if they had a spatial resolution of ≤ 1" or ~0.0083333°. No occurrence data of *M. emarginata* were taken into account, because there are no recent nor precise enough locality data. Similarly, no occurrence data of *M. domingensis* from Haiti were included. This led to a presence-only dataset of 635 occurrences. This dataset was subsequently rarefied with a rarefying distance of 1 km, reducing it to 30 independent occurrences. After rarefying the dataset, too few occurrences remained to perform SDM on each species individually. Instead, we opted for an explicit 'genus-level' approach and performed SDM on all remaining independent occurrences.

Predictor variables were obtained from ENVIREM (Title and Bemmels 2018), CHELSA (Karger et al. 2017) and SEDAC (Venter et al. 2016, 2018). The included variables comprise environmental variables, describing temperature, precipitation, aridity, evapotranspiration and topography, and one variable describing cumulative human pressures on the environment. To identify strongly correlated variables, we used the "raster.cor.matrix" and "raster.cor.plot" function in the R package ENMTools (Warren et al. 2010). The Pearson correlation coefficient "r" was used to determine which variables were most strongly correlated ($|r| > 0.7$). Of the correlated variables, we selected the most ecophysiologically meaningful ones in accordance to Mod et al. (2016) (see Online Resource 4). All variables were obtained with a spatial resolution of 1 km^2 or ~0.0083333°.

Model Fitting and Evaluation

The different pseudo-absence selection strategies explored here resulted in 15 ensemble model predictions depicting the suitability of Hispaniola for the focal *Magnolia* spp. Each ensemble model consisted of eight different modelling algorithms, as implemented in the ensemble modelling framework of the package "sdm" (Naimi and Araújo 2016) in R (R Core Team 2020): two regression-based models (Generalised Linear Model (GLM) & Multivariate Adaptive Regression Splines (MARS)), five machine-learning models (Classification and Regression Trees (CART), Boosted Regression Trees (GBM), Maximum Entropy (MaxEnt), Random Forest (RF) & Support Vector Machine (SVM)) and one likelihood-based model (Maxlike). Each SDM algorithm ran with default settings (Naimi and Araújo 2016). Pseudo-absences were sampled at random in background areas defined as either concentric zones around the presence data (using four different radii (i.e. 10 km, 25 km, 50 km & 75 km)) or across the whole of Hispaniola (i.e. no radius restriction) *sensu* VanDerWal et al. (2009). Pseudo-absences were not selected from pixels containing *Magnolia* presences. Pixels correspond with the spatial resolution of the predictor variables (i.e. 1 km^2). Three prevalence ratios were used to sample pseudo-absences: neutral prevalence (1:1), twice the number of presences (1:2) and five times as much (1:5). Prevalence ratios were kept low to account for a low number of (independent) *Magnolia* occurrences, as proposed by Liu et al. (2019). This led to a total of 15 (5 x 3) predictions of *Magnolia* habitat suitability (on a scale of 0 to 1), which were combined into ensemble predictions. Each individual model was calibrated using a 10-fold cross-validation with 80-20% random split of the presence data to serve as training data for each replicate. Model evaluation is given by the True Skill Statistic (TSS) since it is considered the most optimal measure for the performance of predictive models (Allouche et al. 2006). TSS values range from -1 to +1, with values of zero or less indicating a performance that is no better than random, while a value of +1 suggests a perfect model capacity to discriminate between suitable and unsuitable areas. Interpretation of the TSS values follow Allouche et al. (2006). For each of the 15 ensemble models, only "good performing" algorithms, characterized by a TSS-value of > 0.7, were kept and ensemble habitat suitabilities were obtained using simple averaging of the good performing algorithm outcomes.

Finally, to obtain a single prediction of likely *Magnolia* occurrence across Hispaniola, three different methods were used to summarize the 15 ensemble model outcomes above. First, the 15 *Magnolia* habitat suitability predictions were transformed from continuous suitabilities (0 to 1) to discrete predicted presences versus absences, using the habitat suitability threshold maximizing model TSS as cut-off and setting the suitability for pixels below this threshold to zero. A first final distribution map was then obtained through "vote counting" pixels in every presence-absence prediction, thus

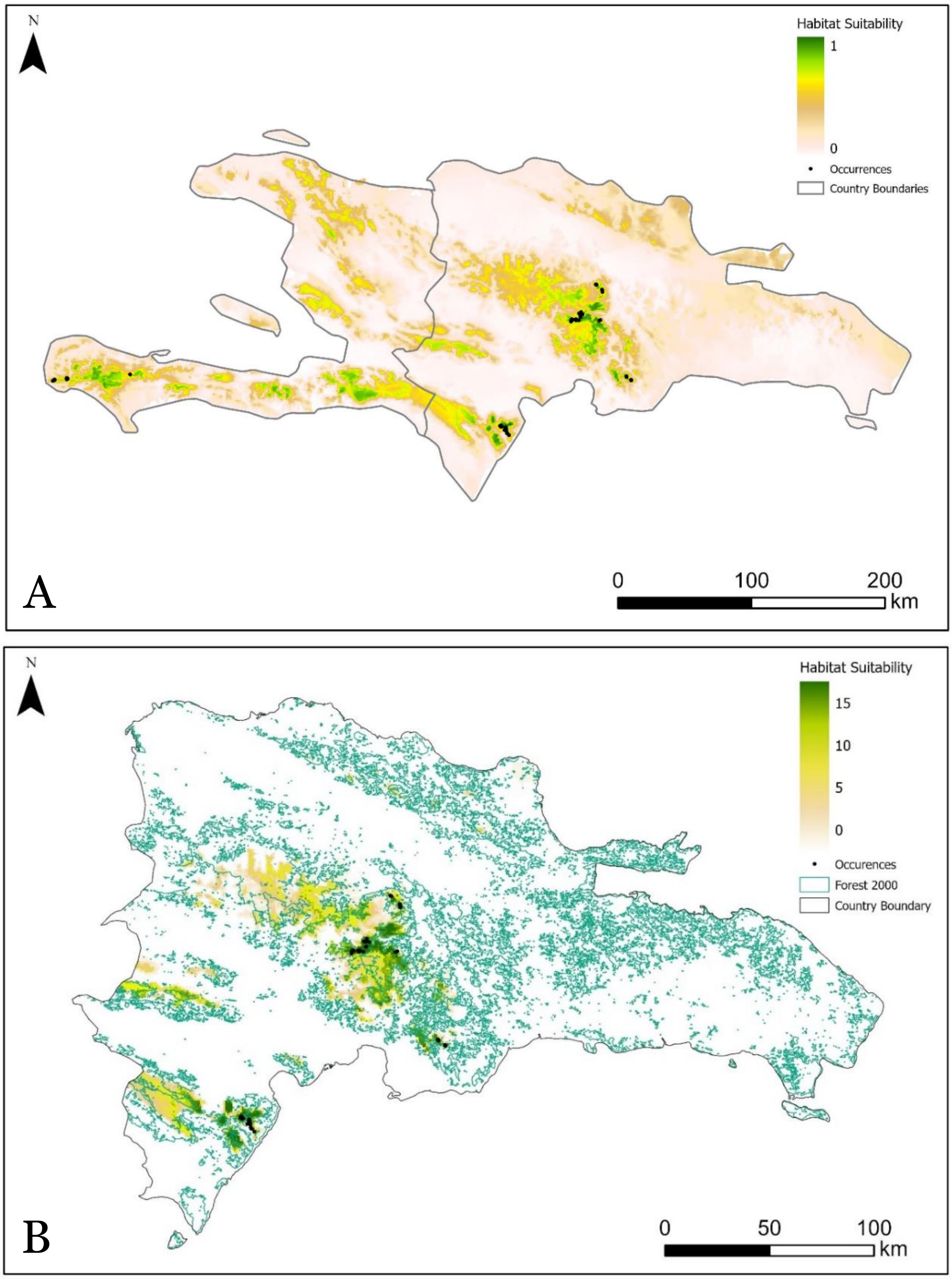


Fig. 3. Ensemble-based distribution map for the *Magnolia* species from Hispaniola and the Dominican Republic, respectively. **A**. ensemble-based distribution map for Hispaniola using weighted averaging. Habitat suitability ranges from zero to one, representing the weighted mean TSS value for each pixel. **B**. ensemble-based distribution map for the Dominican Republic using "vote counting." Habitat suitability ranges from zero to fifteen, representing the number of ensemble analyses that labelled a pixel as suitable (i.e. has a TSS-value > 0.7). Forest 2000 represents the forest cover in the year 2000 (Sangermano et al. 2015).

creating a gradient from 0 (area was never suitable) to 15 (area was always indicated as suitable). Secondly, we used a weighted averaging procedure, combining the 15 habitat suitabilities into a single habitat suitability map, weighing individual models by their TSS values (better models contribute more strongly to the final prediction).

Thirdly, a similar standard "unweighted" averaging was performed. These latter two methods resulted in a continuous prediction of habitat suitability between 0 and 1. These three final maps were further edited to include occurrence data. An additional ensemble distribution map, specifically for the DR, was established from the "vote counting" distribution map of Hispaniola (Fig. 3B). Here, we added an additional layer of forest cover from the year 2000, obtained from Sangermano et al. (2015).

Conservation Genetics

Plant Material and DNA Extraction

Populations of *Magnolia domingensis*, *M. hamorii* and *M. pallescens* in the DR were sampled in 2015 and 2021 by haphazard sampling (Ward and Jasieniuk 2009; Veltjen et al. 2019). In total 417 individuals were genotyped at 12 distinct localities. For the localities sampled in 2015, one herbarium voucher is listed as a morphological voucher. Sample information can be found in Table 1 and a map of the respective sampling locations is given in Fig. 2. Samples were validated using the identification key from Veltjen (2020).

DNA isolation was performed on dried leaf tissue according to a modified cetyltrimethylammonium bromide (CTAB) extraction protocol (Doyle and Doyle 1987), with MagAttract Suspension G solution (Qiagen, Germantown, USA) mediated cleaning (Larridon et al. 2015). A Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, Massachusetts, USA) and Nanodrop 2000 Spectrophotometer (Thermo Fisher Scientific) were used to assess DNA quantity and quality control, respectively.

Microsatellite Testing and Genotypification

Twenty-five microsatellite or Simple Sequence Repeats (SSR) markers from Veltjen et al. (2019) were selected, based on the SSR marker being polymorphic in at least one of the three *Magnolia* species from the DR. From each population of every species, four individuals were selected to re-evaluate the performance of these microsatellite markers (Online Resource 2). The four individuals were chosen to include as much variation as possible on a spatial scale and to have 260/230 and 260/280 optical density (OD) ratios approaching 2. Performance testing was done using simplex PCRs (i.e. one microsatellite marker per PCR) with a three-primer PCR (Vartia et al. 2014). A reaction contained 2 x QIA Multiplex PCR Master Mix (Qiagen), 5ng/µL DNA, 0.025 µM for each forward primer, 0.1 µM for each reverse primer and 0.1 µM for each specified dye. The selected primer pairs were labelled with a

Species	ID	Locality	Loc. ID	RL	Herbarium Reference
M. domingensis	DOM	Loma Barbacoa	BAR	CR	Veltjen et al. 2015-011 (GENT, JBSD)
		Loma Rodríguez	ROD		Veltjen et al. 2015-012 (GENT, HAJB, JBSD)
M. hamorii	HAM	Cachote	CAC	E	Veltjen et al. 2015-010 (GENT, JBSD)
		Cortico	COR		Veltjen et al. 2015-009 (GENT, HAJB, JBSD, K)
		Laguneta	LAG		/
		Tonobán	TON		/
M. pallescens	PAL	Casabito	CAS	E	Veltjen et al. 2015-005 (GENT, JBSD, HAJB)
		Entrance	ENT		Veltjen et al. 2015-006 (GENT, JBSD, HAJB, IEB, K)
		Camino Guayabal	GUA		/
		Loma de la Sal	SAL		Veltjen et al. 2015-004 (GENT, JBSD)
		Montellano	MON		Veltjen et al. 2015-007 (GENT, JBSD)
		Rancho Guaraguao	RAN		/

Table 1. Sample information of the three *Magnolia* species from the Dominican Republic. **Species**: taxonomy according to Howard (1948). **ID**: three letter code to uniquely identify a species. **Loc. ID**: three letter code to uniquely identify a locality. **RL**: Red List status following Rivers et al. (2016); CR: Critically Endangered; E: Endangered. Herbarium acronyms follow the Index Herbariorum (Thiers, [continuously updated]). Sampling was executed in April-May 2015 and January-February 2021.

fluorescent dye FAM, NED, PET and VIC, which were linked to their respective universal tail T3, Hill, Neo and M13. PCRs ran on a volume of 5µL under the following conditions: an initial activation step of 15 min at 95°C, followed by 35 cycles of 30 sec at 94°C, 90 sec at 57°C and 90 sec at 72°C; and a final extension for 10 min at 72°C. An ABI 3730XL fragment analyser (Thermo Fisher Scientific) with a GeneScanTM 500 LIZ™ ladder (Thermo Fisher Scientific) was used to separate and visualise the PCR products. The results of the simplex tests that ran on 24 test individuals were analysed in Geneious v.8.1.9 (http://www.geneious.com, Kearse et al. 2012) using the microsatellite plugin. When a SSR marker could unambiguously be scored for all tested species and populations, and was polymorphic for at least one individual, the microsatellite marker was included in the subsequent multiplex design and final genotyping. Multiplex pools were designed using Multiplex Manager (Holleley and Geerts 2009). PCR conditions and peak calling occurred cfr. the simplex testing above. In total, four multiplex pools allowed unambiguous genotypification of the test-individuals, whereafter we genotyped the full dataset of 417 individuals. In total we evaluated 25 SSR markers of which 17 were considered qualitative and unambiguous.

To account for human and technical flaws such as genotyping errors and null alleles, the final microsatellite dataset was subjected to MICROCHECKER v.2.2.3 (Van Oosterhout et al. 2004) and ML-NullFreq (Kalinowski and Taper 2006). Null alleles are alleles that do not produce a functional end-product, or have a mutation in the primer region which may inhibit PCR amplification (Frankham et al. 2015). Null alleles may lead to preferential amplification of short alleles (i.e. large allele dropout) or slippage during PCR amplification (Gagneux et al. 1997). Therefore, null alleles can obscure genotype frequencies and its derived population statistics. MICROCHECKER and ML-NullFreq ran for 1000 and 100 000 repetitions, respectively. No markers were removed.

To assure random sampling of the genome, we examined the dataset for the presence of linkage disequilibrium (LD) using the program GENEPOP v.4.7.5 (Raymond and Rousset 1995; Rousset 2008). GENEPOP ran with the dememorization number set to 10 000, with 1000 batches and with 50 000 iterations per batch. Following Waples (2015), both uncorrected and (sequential Bonferroni) corrected p-values were considered. One marker, MA40_045, was discarded, bringing the final number of markers to 16 (see Online Resource 3).

Data Analyses

In STRUCTURE v.2.3.4 (Pritchard et al. 2000), analyses were done using six datasets: one with all three focal species, three datasets for each species and one dataset for each genetic cluster of PAL as a result of the first analysis (see further). STRUCTURE ran under the following conditions: 100 000 MCMC replicates after an initial burn-in of 100 000, correlated allelic frequencies and presuming the presence of an admixture model. The number of groups (K) was set from one to fifteen for the general dataset (all species included), from one to ten for the species-specific datasets and from one to five for the genetic clusters from PAL. Each value of K was run ten times for each dataset. A visualization of the results was obtained from Structure Harvester Web v.0.6.94 (Earl and vonHoldt 2012). The optimal K-value was selected based on the ΔK statistic, following Evanno et al. (2005), and the mean maximum likelihood (Mean LnK or $\overline{\text{LnK}}$). Mean LnK proved valuable as ΔK cannot detect single clusters. Visualisation of the STRUCTURE barplots was done with DISTRUCT v.1.1 (Rosenberg 2004). DAPC (Discriminant Analysis of Principal Components) analyses were executed in R using the package "adegenet" (Jombart and Ahmed 2011) on the same datasets as the STRUCTURE analyses. Using the "find.cluster" function, we retained the maximum number of PCs and selected groups on the basis of the Bayesian Information Criterion (BIC). The number of Principal Component Analysis (PCA) eigenvalues was determined using 1000 cross-validation replicates. The PCA of the DAPC was set according to the lowest Mean Squared Error (MSE). All discriminant functions (i.e. DA eigenvalues) were kept. Furthermore, pairwise F_{ST} (Weir and Cockerham 1984), G_{ST} (Nei and Chesser 1983), G'_{ST} (Hedrick 2005) and Jost's D (D_J; Jost 2008) values and their respective confidence intervals were calculated for each population and locality (Table 4, Online Resource 8) using the R package diveRsity (Keenan et al. 2013). Interpretation of the F_{ST}-values follows Hartl and Clark (1997).

Diversity statistics were calculated for each species, population, and locality (Table 2). Genetic diversity was estimated by quantifying the following statistics in GenAlEx 6.503 (Peakall and Smouse 2006, 2012): number of alleles per locus (A), number of private alleles (A_P), observed heterozygosity (H_O), expected heterozygosity (H_E) and the absolute number of polymorphic loci (P). Moreover, allelic richness (A_R) and the inbreeding coefficient (F_{IS}) was calculated in FSTAT v.2.9.4 (Goudet 2000). A_R was calculated for the population and localities based on a minimum sample size of 6, 14 and 36 individuals. To quantify if F_{IS} significantly differed from zero, deviations from Hardy-Weinberg proportions

ID	N_O	A	$A_{R\text{-}6}$	$A_{R\text{-}14}$	$A_{R\text{-}36}$	A_{PP} (A_{PA})	H_O	H_E	F_{IS}	P
DOM										
BAR	36	5.250	3.585	4.412	5.250	25 (4)	0.627	0.587	-0.054	16
ROD	50	4.188	2.972	3.473	3.990	8 (2)	0.538	0.524	-0.016	16
HAM										
BAH	118	8.063	/	/	6.859	46 (-)	0.579	0.605	0.047	15
CAC	52	7.063	4.027	5.318	6.608	11 (7)	0.575	0.575	0.011	13
COR	52	6.625	4.129	5.378	6.329	8 (5)	0.578	0.590	0.029	14
LAG	7	3.750	3.565	/	/	1 (1)	0.616	0.540	-0.064	14
TON	7	4.375	4.153	/	/	4 (3)	0.588	0.574	0.053	14
PAL										
EBV	80	5.813	/	/	5.375	28 (7)	0.546	0.562	0.035*	15
CAS	40	5.125	3.389	4.231	5.039	4 (1)	0.545	0.541	0.004	15
SAL	40	5.438	3.686	4.635	5.376	4 (2)	0.547	0.548	0.016	15
VAL	133	6.250	/	/	5.367	21 (6)	0.527	0.552	0.048	16
ENT	61	5.188	3.349	4.188	4.951	4 (1)	0.520	0.529	0.025*	16
GUA	15	4.063	3.135	4.252	/	2 (1)	0.518	0.482	-0.038	15
MON	41	5.188	3.433	4.003	5.073	3 (3)	0.527	0.537	0.030*	15
RAN	16	4.500	3.496	4.402	/	1 (0)	0.564	0.554	0.015	16

Table 2. Population and locality statistics for the Magnolias of the Dominican Republic. **ID**. Three letter code to uniquely identify a species, population, or locality, following Table 1. **N_O**. Number of sampled individuals. **A**. Average allelic diversity. **$A_{R\text{-}X}$**. Allelic richness based on a minimum sample size of 6, 14 and 36 individuals. **A_P**. Number of private alleles between all localities (**A_{PA}**) or between populations only (**A_{PP}**). **H_O**. Average observed heterozygosity. **H_E**. Average expected heterozygosity. **F_{IS}**. Inbreeding coefficient. Significant (p = 0.05) deviations from Hardy-Weinberg proportions are indicated with an asterisk (*). **P**. Number of polymorphic loci.

(HWP) were tested in GENEPOP with 2-tailed exact tests for every marker x population combination. If possible, a complete enumeration was executed (Louis and Dempster 1987). Otherwise, MCMC chains were employed with 200 batches and 50 000 iteration (Guo and Thompson 1992).

Results

Species Distribution Modelling

Out of 39 variables, 21 were uncorrelated and considered ecophysiologically meaningful and were retained for the subsequent analyses (Online Resource 4). During model evaluation, no ensemble models were discarded. Model performance for each of the 15 ensemble predictions can be found in Table 3. An upward trend of TSS-values is visible for increasing pseudo-absence sampling distance as well as for a higher prevalence ratio. For each of the three methods to summarize the 15 ensemble models (vote counting, weighted and unweighted averaging), maps were made to visualize the habitat suitabilities across Hispaniola and the DR. These maps can be found in Fig. 3 and Online Resource 5 (A-B). Regions of high habitat suitability for the Magnolias of Hispaniola seem to co-occur with the numerous mountain ranges on the islands (Fig 1, Fig.2A).

Microsatellite Markers

Twenty-five microsatellite markers were tested in simplex and multiplex. Seventeen markers were polymorphic for at least one individual and allowed unambiguous genotypification. Subsequently, MICROCHECKER labelled four markers with potential null alleles: MA40_045, MA42_001, MA42_059, MA42_126 and MA42_472. However, they were only highlighted in MON. ML-NullFreq indicated high values (>0.1) for MA42_001 and MA42_059 in locality MON and for MA42_203, MA42_241 and MA42_397 in locality TON. GENEPOP revealed that one marker, MA40_045, showed significant LD with three markers after sequential Bonferroni correction: MA42_231 in ROD, MA40_282 in BAR and finally MA42_472 in ROD, SAL, and MON. Hence, MA40_045 was discarded to guarantee independent sampling of the genome with respect to other microsatellite markers, bringing the final number of markers to a total of 16 (see Online Resource 3).

Population Structure and Genetic Diversity

ΔK and mean LnK plots for all four STRUCTURE-analyses are presented in Online Resource 6. The optimal K derived from the ΔK and mean LnK plots were similar for each analysis except for HAM. For the STRUCTURE analysis of the entire dataset

Prevalence ratio	Pseudo-absence sampling distance				
	10 km	**25 km**	**50 km**	**75 km**	**Random**
1:1	0.830 ± 0.051	0.897 ± 0.053	0.917 ± 0.039	0.950 ± 0.036	0.977 ± 0.022
1:2	0.800 ± 0.069	0.953 ± 0.028	0.943 ± 0.052	0.917 ± 0.032	0.950 ± 0.036
1:5	0.957 ± 0.035	0.950 ± 0.048	0.940 ± 0.041	0.993 ± 0.014	0.993 ± 0.021

Table 3. Model performance, indicated by TSS-values, for the 15 ensemble analyses. TSS-values and their respective standard deviation are rounded to three decimals.

encompassing all species, the optimal K is four. In the species datasets the optimal K for DOM, HAM and PAL is two. The mean LnK plots also depict 2 as the best value for K, except for HAM. Here, the likelihood for K = 1 is not that much lower than K=2 and we consider K=1 to be the optimal number of genetic clusters since ΔK is unable to consider only one cluster. Representative replicate barplots for the optimal K groups are portrayed in Fig. 4. Based on these results, five populations were defined: BAR, ROD, BAH, SAL and VAL. Populations retained their names from Veltjen et al. (2019) if locations resulted in an identical population, otherwise a unique population ID was designated. Fig. 5 visualizes the DAPC analysis on all individuals, showing four genetic clusters in accordance with the STRUCTURE analysis on all species (Fig. 4A). The first axis shows a clear distinction between BAH and the three remaining populations, which are subsequently separated by the second axis. The results from the other DAPC analyses can be found in Online Resource 7. Multiple, unambiguous genetic clusters were obtained for the full dataset and for PAL. Similar to the STRUCTURE-analysis, DOM showed an intermediate result. Pairwise F_{ST}, G_{ST}, G'_{ST} and Jost's D (D_J) values can be found in Table 4 and Online Resource 8. Values between populations range from 0.081 to 0.271 for F_{ST}, from 0.042 to 0.158 for G_{ST}, from 0.149 to 0.550 for G'_{ST} and from 0.081 to 0.378 for D_J. Confidence intervals (CIs) of the pairwise indices can be found in Online Resource 9.

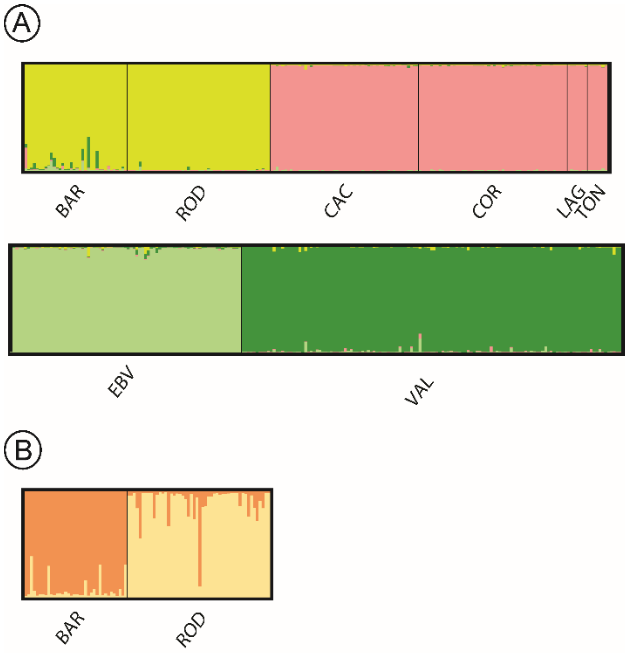


Fig. 4. STRUCTURE barplots for the Magnolias of the Dominican Republic and *Magnolia domingensis* specifically. Optimal K-values are given based on mean LnK. **A**. Dataset comprising all three species, K = 4. **B**. *Magnolia domingensis*, K = 2.

Diversity statistics for each population, locality and microsatellite marker are compiled in Online Resource 10. The mean values for the individual populations and localities can be found in Table 2. One population and two localities, namely EBV, ENT and MON, showed significant deviations from HWP for seven, four and one out of 16 markers, respectively, which resulted in an F_{IS} that was significantly deviating from zero.

Discussion

Evaluation of Species Distribution Modelling and Microsatellites

The ensemble SDM predicted the habitat of *Magnolia* across Hispaniola with supposedly great accuracy. This is visible in Table 3, which indicates the model performance across prevalence ratios and pseudo-absence sampling distance. These values range from 0.800 to 0.993, which are promising results as values greater than 0.7 are indicative for good predictive models (Allouche et al. 2006). Furthermore, all 66 additional occurrences from the sampling effort in 2021 fell within highly suitable predicted areas. Similarly, herbarium occurrence data that was deemed too unprecise for the SDM analyses, were often situated in highly suitable habitats (green to dark green areas in Fig. 3A). This validation of predictive power of the SDMs has to be nuanced since all new sampled locations were in relatively close proximity to known populations from

◂**Table 4.** Pairwise fixation and allelic differentiation statistics (F_{ST}, G_{ST}, G'_{ST} & D_J) for every population combination.

Genetic differentiation statistics					
Statistic	Pop	BAR	ROD	BAH	SAL
F_{ST}	ROD	0.081			
G_{ST}		0.042			
G'_{ST}		0.149			
D_J		0.081			
F_{ST}	BAH	0.208	0.257		
G_{ST}		0.117	0.151		
G'_{ST}		0.468	0.550		
D_J		0.372	0.378		
F_{ST}	SAL	0.215	0.271	0.219	
G_{ST}		0.120	0.158	0.124	
G'_{ST}		0.450	0.540	0.475	
D_J		0.268	0.328	0.369	
F_{ST}	VAL	0.216	0.264	0.218	0.147
G_{ST}		0.119	0.153	0.122	0.079
G'_{ST}		0.439	0.515	0.460	0.280
D_J		0.281	0.302	0.352	0.154

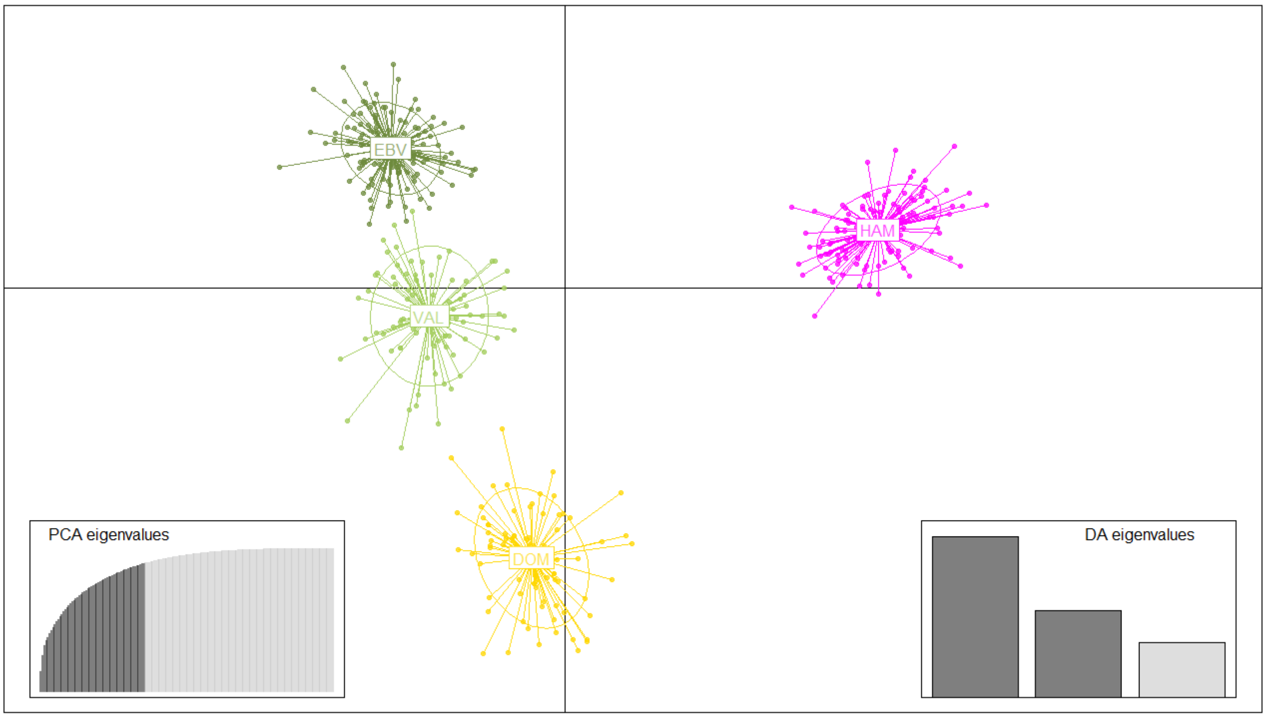


Fig. 5. DAPC (Discriminant Analysis of Principal Components) plots for the Magnolias of the Dominican Republic. Names follow the species ID (if K = 1) or population ID (if K > 1), depending on the number of clusters (K). Four clusters are visible across three species: DOM, HAM, EBV and VAL. The latter two belong to PAL.

the DR. Moreover, high TSS values may indicate model overfitting as model evaluation metrics are both a function of prediction accuracy as well as modelling conditions (Woodcock 1976; Wunderlich et al. 2019).

During microsatellite testing, a downward trend of qualitative, usable markers was observed (Hodel et al. 2016). Starting with 25 markers from Veltjen et al. (2019) which ought to be polymorphic, 17 proved to be qualitative and unambiguous after genotypification. Subsequently, tests on null alleles indicated several markers to contain potential null alleles in MON and TON (Table 1). However, as TON is a locality with a small sample size, we do not consider these results as a true indication for null alleles. Similarly, as null alleles were only found in MON and in no other locality, no markers were removed from the subsequent downstream analyses. The presence of LD between the locus pairs MA40_045 – MA40_282, MA40_045 – MA42_231 and MA40_045 – MA42_472, led to the subsequent removal of marker MA40_045. This reduced the final number of microsatellite markers to 16. Interestingly, due to an increment in sample size compared to Veltjen et al. (2019), we no longer see LD for population ROD. In Veltjen et al. (2019), this was attributed to a recent bottleneck and a subsequent lack of random mating to restore linked loci (Slatkin 2008). Now, we can elucidate that this was a random sampling error, but see Waples (2015).

Guiding Explorations Using the Modelled Suitable Habitat

Mountain ranges with the most promising areas of detecting new *Magnolia* populations in the DR are the Cordillera Central, the Sierra de Bahoruco and the Sierra de Neiba. Possible remnant populations could be found in the Cordillera Septentrional and the Sierra Martín García, an isolated extension of the Sierra de Neiba. In the DR, areas with high habitat suitability mainly reside in areas of the Forest 2000 layer (Fig. 3B). For Haiti, areas of interest reside in the mountain ranges Massif de la Hotte and Massif de la Selle, the Haitian counterpart of the Sierra de Bahoruco. To a lesser degree, possible areas of interest can be found in Massif du Nord, Montagnes Noires and Chaîne des Matheux (Fig. 3A). In Haiti, expeditions to Massif de la Selle have did not find any *Magnolia* species (Pers. comm. Joel Timyan, Société Audubon Haiti). Due to an extensive history of deforestation and the absence of dark green patches of suitable habitat in northern Haiti (Fig. 3A), new populations in mountain ranges aside from Massif de la Hotte and Massif de la Selle are deemed unlikely (Hedges et al. 2018). Knowledge about the potential habitat provides policy makers and on-the-ground conservationists with a guideline for spatial prioritisation of conservation efforts (Villero et al. 2017). Concretely, it can serve as the basis for reinforcement efforts and the designation of habitat corridors; and allows to search for the true or full extent of populations more efficiently (Williams et al. 2009). Habitat corridors have successfully been applied in Baruah et al. (2019), Adhikari et al. (2012) and Liu et al. (2018).

Defining Conservation Units

By integrating the results of the SDM (Fig. 3, Online Resource 5) and conservation genetic analyses (Fig. 4 – 5, Table 2 & 4, Online Resource 6 – 9), we propose to treat the five genetic *Magnolia* populations of the DR (Fig. 4) as three conservation units (CU), following the morphospecies as described by Howard (1984) and confirmed with recent *in situ* observations (Castillo et al. 2018; Veltjen 2020). Conservation or management units are groups of individuals that are clustered together to contain the genetic diversity necessary to ensure evolvability in the light of changing environments and local adaptation (Weckworth et al. 2018). In the following three paragraphs we deliver the argumentation species-by species to treat them as one CU each.

According to the STRUCTURE and DAPC analysis (Fig. 4 – 5, Online Resource 7D - F), *M. pallescens* is divided in two populations: a northern population (EBV) in Ébano Verde Scientific Reserve and a southern population (VAL) in Valle Nuevo National Park (Fig. 3B). Furthermore, there are no recent observations of *Magnolia* in the suitable habitat patch between EBV and VAL (Fig. 2, Fig. 3B). Hence, we see distinct spatial and genetic grouping. Three herbarium vouchers (Garcia_1002, Garcia_1184 and Veloz_3237) do imply a historical distribution of *M. pallescens* in which the two populations were connected. These vouchers were excluded from the SDM analysis. The division of populations is reinforced by the pairwise genetic differentiation statistics from Table 4. Between EBV and VAL, we see an F_{ST}-value of 0.147, which is considered moderate genetic differentiation following Hartl and Clark (1997). Furthermore, a value of 0.154 for D_J suggests sufficient genetic differentiation to be considered independent populations, especially when compared to D_J-values between population of different species, ranging from 0.268 to 0.378. When we compare infraspecific F_{ST}- and D_J-values with other studies on island tree endemic populations, we find similar results. For example, F_{ST} = 0.000 – 0.229 for *Pinus caribaea* var. *bahamensis*, an endemic variety from the Bahaman archipelago (Sanchez et al. 2014). In terms of genetic diversity, the population summary statistics for EBV and VAL indicate average genetic health when compared with other *Magnolia* populations from the DR (Table 2). Notably, EBV is the only population to exhibit a significant, but low, F_{IS}-value. Despite the apparent genetic substructure of the species, we propose to treat *M. pallescens* as one CU due to the significant F_{IS}-value of population EBV and localities ENT and MON (population VAL); and the potential historical connection between the populations.

For *M. domingensis*, it is less straightforward to state the number of (genetic) populations, with both arguments in favor of one and two populations. Although the STRUCTURE (Fig. 4A) and DAPC (Fig. 5) analyses on the full dataset depict *M. domingensis* individuals to be one genetic cluster, the species-specific analyses (Fig. 4B & 5, Online Resource 7B) retrieve two with a few migrants/relict trees in each of the localities. Considering the habitat suitability of the area, connecting the two populations, there is no strong indication for any spatial schisma. The two populations are connected by a continuous mountain ridge and only a minor reduction in habitat suitability of the connecting pixels is visible. This raises the question whether *M. domingensis* effectively consists of two populations or whether this distinction is artificial. Taking into account the genetic differentiation indices (Table 4), we see F_{ST}- and D_J-values between BAR and ROD (F_{ST}: 0.081; D_J: 0.081) that reside in the same category (moderate genetic differentiation; Hartl and Clark 1997) as the values between EBV and VAL (F_{ST}: 0.147; D_J: 0.154), but in the minimum as opposed to the maximum values of the range for that category (i.e. 0.05 – 0.15). These intraspecific values are comparable with a study on *Metrosideros* species from Hawaii (F_{ST} = 0.09 – 0.12; Harbaugh et al. 2009). If we look at the indices between localities within the populations of Magnolia in the DR (Online Resource 8), values range from 0.023 to 0.078 for F_{ST} and from 0.011 to 0.074 for D_J. As these maximum values closely approximate the values between BAR and ROD, we are inclined to conclude that BAR and ROD should be considered localities within one population. However, these maxima are derived from values between localities with a small number of sampled individuals (GUA, LAG, RAN and TON; Table 1 – 2) that potentially split one large continuous population. Considering only populations with N > 39, the maximum range of F_{ST} and D_J drops down to 0.049 and 0.068, respectively. The high number of private alleles (A_{PP}; Table 2) in population BAR compared to ROD confirm this apparent differentiation. To summarize, the STRUCTURE (Fig. 4B) and DAPC (Online Resource 7B) analyses on the species-specific dataset indicate two populations, while these analyses on the full dataset (Fig. 4A, Fig. 5) and the SDM map (Fig. 3) only predict one. Additionally, the genetic differentiation statistics indicate intermediate values (Table 4) and there is a notable number of private alleles that sets these localities apart (Table 2). Although the signal is not as strong as the *M. pallescens* populations, there is enough genetic evidence to state that we have two recently differentiated populations. The genetic diversity of BAR is average across the Magnolias of the DR (Table 2). In contrast, ROD exhibits the worst genetic health among the three studied species, with the lowest values for any A-statistic. However, both populations have low and non-significant F_{IS}-values. As ROD is genetically degraded, genetic substructuring is not pronounced and the localities are linked by suitable habitat, we consider *M. domingensis* to be one CU. This approach allows genetic rescue effects between these populations to negate a further decline in genetic diversity (Ingvarsson 2001; Whiteley et al. 2014).

By integrating the genetic structuring results of *M. ekmanii* from Veltjen et al. (2019) with the SDM analyses (Fig. 3A), we see a similar issue as the spatial segregation of *M. domingensis*. The species is split in two populations, Morne Grand Bois (GRA) and Morne Mansinte (MAN), evidenced by a STRUCTURE and DAPC analysis. Based on our SDM results (Fig. 3A), this disjunction is unexpected considering that both populations are connected by a stretch of mountains with highly suitable potential habitat. Further sampling efforts in the area ought to resolve this issue. Considering its genetic substructure and troublesome genetic health, we treat *M. ekmanii* as one CU. Noteworthy, new occurrences of *M. ekmanii* were found relatively recently in the region of Bois Pangnol (easternmost point on Fig. 2), an isolated location given the geographic distance between other Magnolias in Hispaniola. As these were not sampled for any genetic analysis, we cannot make inferences with regard to their genetic structure or health. We recommend including this new locality in future conservation genetic studies.

M. hamorii only has one population, BAH, situated in the Natural Protected Landscape Miguel Domingo Fuerte. It is the most differentiated from any other population from the DR with respect to D_J (Table 4). When considering the DAPC plot (Fig. 5) and the ensemble-based distribution map (Fig. 3), *M. hamorii* is situated the furthest from any other population from the DR, both in terms of kilometres and genetic distance. This is also visible in the genetic differentiation indices of Table 4. Many of the genetic diversity indices (A, H_E) of *M. hamorii* are the highest among the Magnolias of the DR (Table 2). The high number of private alleles (A_{PP}) further shows that *M. hamorii* is very distinct from the other Magnolias of the DR. Its F_{IS}-value of 0.047 is the highest among all populations, but non-significant and unproblematic as such. We consider *M. hamorii* as one CU since the sampled localities have one continuous modelled suitable habitat (Fig. 3B), they show low genetic differentiation (Table 4) and hence, there is no detectable genetic substructure (Fig. 4A & 5).

Priorities Based on Genetic Diversity

Because no non-threatened, related *Magnolia* species were included in this research, we can only make statements about the genetic diversity being healthy, low or high when comparing the localities and populations with each other (Spielman et al. 2004; Väli et al. 2008). Comparisons with diversity statistics generated in other studies can be made but should be treated with caution. We will discuss these statistics in view of the previously instated population structure, the habitat availability and the species' Red List assessment of Rivers et al. (2016).

Solely comparing the population statistics for the five populations, we do not find dramatic results in terms of genetic health. In general, only three significant F_{IS}-values are found across all populations and localities: population EBV and two localities within population VAL (ENT and MON). Even so, these values are relatively low compared to other *Magnolia* species (e.g. *M. nuevoleonensis* – population SPE and PE: F_{IS} = 0.583* and 0.749* (Chávez-Cortázar et al. 2021); *M. stellata* – population Asahidani: F_{IS} = 0.233* (Tamaki et al. 2016)). Both populations, EBV and VAL, have average values for all population statistics, except for a significant deviation from HWP of the F_{IS}-value for EBV. Peculiarly, EBV has an average number of individuals compared to other DR *Magnolia* species (Table 2), yet it is the only population with a significant F_{IS}-value. Interestingly, no significant F_{IS}-values are found for the two localities in EBV. One potential explanation could be that a Wahlund-effect was artificially created by merging the two localities (De Meeûs 2018). The Wahlund effect reduces heterozygosity in a population due to population substructuring (Wahlund 1928). The opposite situation as EBV is visible in VAL. We attribute this to unrepresentative sampling across a small spatial scale. When comparing these results with Veltjen et al. (2019), no drastic differences are visible, aside from the significant F_{IS}-value for EBV. *Magnolia domingensis,* shows the highest H_O of all populations in BAR, but no other statistic stands out. ROD, however, has the lowest A-, A_R- and A_P-values of all. Despite this, F_{IS}-values remain low and non-significant for both populations. We see that low values for the A statistics in ROD could explain the higher values of genetic differentiation (Table 4) and the STRUCTURE and DAPC results (Fig. 4, Online Resource 7B). *Magnolia hamorii* with its sole population BAH has the highest reported values for A, A_R, A_{PP} and H_E. The pattern of limited gene flow, leading to high genetic structuring, and little inbreeding is similar to other studies such as Aldaba Núñez et al. (2021) and Veltjen et al. (2019). Inbreeding is expected to be minimal as the flowers of *Magnolia* are protogynous (Thien 1974) and promote outcrossing (Tamaki et al. 2009), but not absent as geitonogamy has been reported in other *Magnolia* species (Bernhardt and Thien 1987; Hirayama et al. 2005; Tamaki et al. 2008). The strong inbreeding depression in the early life stages of trees ensures that adult plants result solely from outcrossing (Sorensen 1999), increasing outcrossing rates and maintaining genetic diversity (Petit and Hampe 2006). High genetic diversity is desirable as their sessile lifestyle allows evolution of locally adapted ecotypes. Additionally, genetic diversity is "stored" in their great longevity, allowing genetic variants to circulate a long time within the populations (Alberto et al. 2013). This allows for potential reinforcement of genetically degraded populations (Aitken et al. 2008).

Unexpected results are found when comparing the species' genetic diversity with their habitat availability and Red List assessment category of Rivers et al. (2016). *Magnolia domingensis* is Critically Endangered (CR; Wheeler 2015), has the worst genetic health of the Magnolias of the DR and has the least amount of suitable *Magnolia* habitat in its near vicinity (Fig. 3B), being restricted to the Padre Luis Quinn National Park. *Magnolia hamorii* is Endangered (E; Global Tree Specialist Group 2014), has more suitable habitat in its vicinity than *M. domingensis*, and despite it being reduced to one population, has the highest genetic diversity statistics. *Magnolia pallescens*, Endangered (E; Global Tree Specialist Group 2014b) with the most available suitable habitat of all the magnolias of Hispaniola by far, has only a mediocre genetic health. When we consider the population statistics for *M. ekmanii* (CR; Wheeler and Rivers 2020) of Veltjen et al. (2019) together with the SDM analysis (Fig. 3A), *M. ekmanii* has a relatively large potential habitat in Massif de la Hotte, exceeding those of *M. domingensis* and *M. hamorii*, but also the worst population statistics among the Magnolias of Hispaniola. We do not see a clear correlation between available (potential) habitat and Red List category or genetic diversity, nor between the latter two. This is surprising as geographic range was the main criteria for *M. hamorii* and *M. pallescens* to be labelled as "Endangered." This mismatch has already been reported in publications such as Vellend and Geber (2005) and Rivers et al. (2014).

Conservation Recommendations

To ensure a sustainable future for the Magnolias of Hispaniola, we propose the following conservation strategy. We prioritize *M. domingensis* and *M. ekmanii* for conservation. *Magnolia ekmanii* because its low genetic diversity was demonstrated in Veltjen et al. (2019) and because its distribution is situated in an region susceptible to deforestation (Hedges et al. 2018). Similarly, *M. domingensis* is prioritized because of the degrading genetic health of population ROD and since both populations are distributed in an isolated extension of the Cordillera Central. *Magnolia pallescens* was not prioritized because the inbreeding in EBV can easily be reduced with propagation efforts and because it has sufficient suitable *Magnolia* habitat in its vicinity (Fig. 3B). For the Magnolias of Hispaniola, the most effective conservation actions that we propose are exploration, protection, and reinforcement.

Exploration aims to further elucidate the contemporary distribution of the Magnolias of Hispaniola using Fig. 3 as a guideline. We propose the following strategy to search for new *Magnolia* individuals/populations, with decreasing priority. Initially, the highly suitable areas between the populations of *M. domingensis*, *M. pallescens* and *M. ekmanii* should be explored, followed by nearby highly suitable, mountainous areas in Massif de la Hotte, Sierra de Bahoruco and Cordillera Central. Subsequently, search for remnant populations of *M. emarginata* and/or *M. domingensis* in Massif du Nord and Montagnes Noires. Finally, areas with no prior indication of *Magnolia* populations such as Sierra de Neiba, Sierra Martín Garcia and Chaîne des Matheux should be further explored. Newfound individuals may lead to the formation or expansion of (new) protected areas (i.e. National Parks and Scientific Reserves) as these charismatic trees are likely candidates to serve as flagship and/or umbrella species (Veltjen 2020).

Protection aims to conserve and expand the species' habitat, TMCFs. Given that an adequate amount of genetic diversity is available within the conservation units, the focus should lie on the expansion of protected areas. Ideally, protected areas should strive to connect forest fragments in the landscape and the different populations as such. These habitat corridors (Christie and Knowles 2015) are especially important for the populations of *M. ekmanii* and *M. domingensis*, should no natural corridor exist. When successful, gene flow between populations should increase and enhance genetic diversity. For *M. hamorii*, protection of habitat is a crucial step in ensuring its survival.

Although protection is considered the most long-term important goal, it often takes time to implement. Hence, we propose to immediately apply reinforcement to bridge this temporal gap. Reinforcement aims to preserve and enhance genetic diversity and population sizes. Our interpretation of "reinforcement" follows Seddon et al. (2014): "The release of an organism into an existing population of conspecifics to enhance population viability." Since gene flow between populations is limited, illustrated by the lower genetic health of ROD and a significant inbreeding value in EBV, we propose to execute reinforcement efforts between populations (Brichieri-Colombi and Moehrenschlager 2016). Here, populations BAR and VAL will function as source populations of genetic diversity and ROD and EBV as their respective receptive sink populations. The benefits of minimizing loss of genetic diversity due to inbreeding and fragmentation outweighs the undoing of potential local adaptations. We encourage conservation practitioners to raise seedlings in nurseries and plant them close/within the sink population. Moreover, new populations could be established in suitable *Magnolia* habitat in the vicinity of known populations. Preferably, in an area secure from large scale threats. Nurseries should avoid hybridization and potential subsequent outbreeding depression (Keller et al. 2000; Tamaki et al. 2017) at any cost. In the case of *M. pallescens*, we propose the suitable area in between the existing populations (Fig. 3B). Propagation efforts in nurseries are ongoing for the Magnolias of the DR, with mixed results (Castillo et al. 2018). We emphasize the importance of collecting seeds from multiple localities and from as many different trees as possible across a large spatial scale. This ensures representative subsampling of the genetic diversity within the conservation unit (Oldfield and Newton 2012). To ensure success, a monitoring programme is to be set up to keep track of the survival rates of the seedlings and mature trees. This allows a quick and swift respond to new threats (Godefroid et al. 2011).

Conclusions

The SDM analyses suggests that the distribution of potential habitats is restricted to the various mountain ranges on Hispaniola. We indicate the mountains of the Montagnes Noires as a priority area to search for the potentially extinct *M. emarginata* and *M. domingensis* in Haiti. The Sierra de Neiba is the most promising area, without prior known *Magnolia* occurrences, to search for new populations in the Dominican Republic. Highly suitable habitat is present between populations of *M. domingensis, M. ekmanii* and *M. pallescens*. These patches are highlighted for future reinforcement efforts should no natural habitat corridor exist. The Magnolias of the DR consist of five genetic populations, representing three species. We recommend each species to be treated as its own conservation units (CU). The pattern of little inbreeding for Caribbean *Magnolia* species seems to persist in our sampling. Population Ébano Verde and localities Entrance and Montellano exhibit minimal, but significant, inbreeding. Out of the three DR *Magnolia* species, *M. domingensis* is highlighted as it contained the lowest genetic diversity, the least number of (known) individuals and the lowest number of sampling localities. For conservation management, we propose to focus on three actions: exploration, protection, and reinforcement. The main recommendation includes protection and/or instatement of habitat corridors between populations of *M. domingensis* and *M. ekmanii*, whereby the SDM results will serve as a guideline for spatial prioritization.

Acknowledgements

We are grateful to the authorities of the Dominican Republic and Haiti for issuing the permits DQCSPAP/SIC/365 and VAPB-01488, respectively. We are grateful to Andy Vierstraete, Pieter Asselman and Ana María Rueda Urrego for their help in the lab. Also, we thank Rosa Rodríguez Peña, Majela Hernández Rodríguez, Víctor Martín Vélez, Yommy Piña, Jean-Francois Orilién Beauduy, William Cinéa, Joel Timyan and Roland Trézil for their help with the fieldwork. We thank Thibeaut Formesyn for his help with the GIS map production. A special thanks to everybody from the Systematic and Evolutionary Botany and Mycology research groups at Ghent University for contributing to this amazing workspace.

Declarations

Funding

We thank Fondation Franklinia, Fundación PROGRESSIO, Ghent University (BOF.DOC.2014.0062.01) and the Research Foundation Flanders (V4.453.15N) for their financial support.

Conflict of interest/Competing interests

The authors hereby declare no conflict of interests.

Availability of data and material

The authors confirm that the supporting data of these findings are available within the article and its supplementary materials.

Code availability

Not applicable.

Authors' contributions

Conceptualization: Emily Veltjen, Marie-Stéphanie Samain; Methodology: Diederik Strubbe, Emily Veltjen; Formal analysis and investigation: Tim Claerhout; Writing – original draft preparation: Tim Claerhout; Writing – review and editing: Diederik Strubbe, Emily Veltjen, Lars Chatrou, Marie-Stéphanie Samain, Ramón Elías Castillo Torres, Tim Claerhout; Funding acquisition: Emily Veltjen, Marie-Stéphanie Samain; Resources: Emily Veltjen, Ramón Elías Castillo Torres; Supervision: Diederik Strubbe, Emily Veltjen, Lars Chatrou.

Ethics approval

Not applicable.

Consent to participate

All authors hereby consent to participate.

Consent for publication

All authors hereby consent for publication.

References

Adhikari D, Barik SK, Upadhaya K (2012) Habitat distribution modelling for reintroduction of *Ilex khasiana* Purk., a critically endangered tree species of northeastern India. Ecol Eng 40:37–43. https://doi.org/10.1016/j.ecoleng.2011.12.004

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

Alberto FJ, Aitken SN, Alía R, et al (2013) Potential for evolutionary responses to climate change - evidence from tree populations. Glob Chang Biol 19:1645–1661

Aldaba Núñez FA, Veltjen E, Martínez Salas EM, Samain MS (2021) Disentangling species delineation and guiding conservation of endangered Magnolias in Veracruz, Mexico. Plants 10:673. https://doi.org/10.3390/plants10040673

Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J Appl Ecol 43:1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x

Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: How, where and how many? Methods Ecol Evol 3:327–338. https://doi.org/10.1111/j.2041-210X.2011.00172.x

Baruah PS, Deka K, Lahkar L, et al (2019) Habitat distribution modelling and reinforcement of *Elaeocarpus serratus* L. - A threatened tree species of Assam, India for improvement of its conservation status. Acta Ecol Sin 39:42–49. https://doi.org/10.1016/j.chnaes.2018.06.002

Bernhardt P, Thien LB (1987) Self-isolation and insect pollination in the primitive angiosperms: New evaluations of older hypotheses. Plant Syst Evol 156:159–176. https://doi.org/10.1007/BF00936071

Brichieri-Colombi TA, Moehrenschlager A (2016) Alignment of threat, effort, and perceived success in North American conservation translocations. Conserv Biol 30:1159–1172

Bubb P, May I, Miles L, Sayer J (2004) Cloud Forest Agenda. UNEP-WCMC, Cambridge, UK

Bystriakova N, Peregrym M, Erkens RHJ, et al (2012) Sampling bias in geographic and environmental space and its effect on the predictive power of species distribution models. Syst Biodivers 10:305–315

Cano-Ortiz A, Musarella CM, Fuentes JCP, et al (2016) Distribution patterns of endemic flora to define hotspots on Hispaniola. Syst Biodivers 14:261–275. https://doi.org/10.1080/14772000.2015.1135195

Castillo RE (2016) Magnolia Society - 3rd International Magnolia Symposium - Cuba. In: Magnolia Soc. https://www.magnoliasociety.org/page-1813213. Accessed 3 Mar 2020

Castillo RE, Encarnación Y, Peguero B, et al (2018) Plan de acción de conservación integrada de las Magnolias (Magnoliaceae) amenazadas de República Dominicana – *Magnolia domingensis*, *M. hamorii* y *M. pallescens*. Fundación PROGRESSIO y Jardín Botánico Nacional Dr. Rafael M. Moscoso, República Dominicana

Chávez-Cortázar A, Oyama K, Ochoa-Zavala M, et al (2021) Conservation genetics of relict tropical species of *Magnolia* (section Macrophylla). Conserv Genet 22:259–273. https://doi.org/10.1007/s10592-021-01334-5

Christie MR, Knowles LL (2015) Habitat corridors facilitate genetic resilience irrespective of species dispersal abilities or population sizes. Evol Appl 8:454–463. https://doi.org/10.1111/eva.12255

Cires E, De Smet Y, Cuesta C, et al (2013) Gap analyses to support ex situ conservation of genetic diversity in *Magnolia*, a flagship group. Biodivers Conserv 22:567–590

Damaris Zurell (2020) Introduction to species distribution modelling (SDM) in R. https://damariszurell.github.io/SDM-Intro/. Accessed 24 Mar 2020

De Meeûs T (2018) Revisiting FIS, FST, wahlund effects, and null alleles. J Hered 109:446–456. https://doi.org/10.1093/jhered/esx106

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361. https://doi.org/10.1007/s12686-011-9548-7

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fajardo S, García-Galvan, R. F, et al (2016) Advances in the knowledge of the vegetation of Hispaniola (Caribbean Central America). Intech i:13. https://doi.org/http://dx.doi.org/10.5772/57353

Figlar RB, Nooteboom HP (2004) Notes on Magnoliaceae IV. Blumea 49(1):87–100

Frankham R, Ballou JD, Briscoe DA (2015) Introduction to Conservation Genetics, 2nd edn. Cambridge University Press, Cambridge, UK

Gagneux P, Boesch C, Woodruff DS (1997) Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. Mol Ecol 6:861–868. https://doi.org/10.1111/j.1365-294X.1997.tb00140.x

Global Tree Specialist Group (2014a) *Magnolia hamorii* (Caimoni). In: IUCN Red List Threat Species 2014 eT193950A2291933. https://dx.doi.org/10.2305/IUCN.UK.2014-1.RLTS.T193950A2291933.en. Accessed 20 May 2021

Global Tree Specialist Group (2014b) *Magnolia pallescens*. In: IUCN Red List Threat Species 2014 eT193985A2293382. https://dx.doi.org/10.2305/IUCN.UK.2014-1.RLTS.T193985A2293382.en. Accessed 20 May 2021

Godefroid S, Piazza C, Rossi G, et al (2011) How successful are plant species reintroductions? Biol Conserv 144:672–682

Goudet J (2000) FSTAT (Version 1.2): A computer program to calculate F-statistics. J Hered 91:509–511. https://doi.org/10.1093/jhered/91.6.509

Guillera-Arroita G, Lahoz-Monfort JJ, Elith J, et al (2015) Is my species distribution model fit for purpose? Matching data and models to applications. Glob Ecol Biogeogr 24:276–292. https://doi.org/10.1111/geb.12268

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8:993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:372. https://doi.org/10.2307/2532296

Hamilton LS, Juvik JO, Scatena FN (1995a) The Puerto Rico tropical cloud forest symposium: introduction and workshop synthesis. In: Tropical montane cloud forests ecological studies, 1st edn. Springer, New York, NY, pp 1–23

Hamilton LS, Juvik JO, Scatena FN (1995b) Tropical montane cloud forests ecological studies, 1st edn. Springer, New York, NY

Hao T, Elith J, Lahoz-Monfort JJ, Guillera-Arroita G (2020) Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. Ecography 43:549–558. https://doi.org/10.1111/ecog.04890

Harbaugh DT, Wagner WL, Percy DM, et al (2009) Genetic structure of the polymorphic *Metrosideros* (Myrtaceae) complex in the Hawaiian Islands using nuclear microsatellite data. PLoS One 4:e4698. https://doi.org/10.1371/journal.pone.0004698

Hartl D, Clark A (1997) Principles of population genetics., 3rd edn. Sinauer, Sunderland, USA

Hedges SB, Cohen WB, Timyan J, Yang Z (2018) Haiti's biodiversity threatened by nearly complete loss of primary forest. Proc Natl Acad Sci USA 115:11850–11855. https://doi.org/10.1073/pnas.1809753115

Hedrick PW (2005) A standardized genetic differentiation measure. Evolution (N Y) 59:1633–1638. https://doi.org/10.1111/j.0014-3820.2005.tb01814.x

Hirayama K, Ishida K, Tomaru N (2005) Effects of pollen shortage and self-pollination on seed production of an endangered tree, *Magnolia stellata*. Ann Bot 95:1009–1015. https://doi.org/10.1093/aob/mci107

Hodel RGJ, Segovia-Salcedo MC, Landis JB, et al (2016) The report of my death was an exaggeration: a review for researchers using microsatellites in the 21st century. Appl Plant Sci 4:1600025. https://doi.org/10.3732/apps.1600025

Holleley CE, Geerts PG (2009) Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. Biotechniques 46:511–517. https://doi.org/10.2144/000113156

Howard RA (1948) The morphology and systematics of the West Indian Magnoliaceae. Bull Torre Bot Club 75(4):335–357

Ingvarsson PK (2001) Restoration of genetic variation lost - the genetic rescue hypothesis. Trends Ecol Evol 16:62–63

IUCN - SSC Species Conservation Planning Sub-Committee (2017) Guidelines for species conservation planning, version 1. IUCN, Gland, Switzerland

Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Jost L (2008) GST and its relatives do not measure differentiation. Mol Ecol 17:4015–4026. https://doi.org/10.1111/j.1365-294X.2008.03887.x

Kalinowski ST, Taper ML (2006) Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. Conserv Genet 7:991–995. https://doi.org/10.1007/s10592-006-9134-9

Karger DN, Conrad O, Böhner J, et al (2017) Climatologies at high resolution for the earth's land surface areas. Sci Data 4:1–20. https://doi.org/10.1038/sdata.2017.122

Kearse M, Moir R, Wilson A, et al (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Keenan K, Mcginnity P, Cross TF, et al (2013) diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol 4:782–788. https://doi.org/10.1111/2041-210X.12067

Keller M, Kollmann J, Edwards PJ (2000) Genetic introgression from distant provenances reduces fitness in local weed populations. J Appl Ecol 37:647–659. https://doi.org/10.1046/j.1365-2664.2000.00517.x

Kramer AT, Havens K (2009) Plant conservation genetics in a changing world. Trends Plant Sci 14:599–607

Larridon I, Walter HE, Guerrero PC, et al (2015) An integrative approach to understanding the evolution and diversity of *Copiapoa* (Cactaceae), a threatened endemic Chilean genus from the Atacama desert. Am J Bot 102:1506–1520. https://doi.org/10.3732/ajb.1500168

Liu C, Newell G, White M (2019) The effect of sample size on the accuracy of species distribution models: considering both presences and pseudo-absences or background sites. Ecography 42:535–548. https://doi.org/10.1111/ecog.03188

Liu C, Newell G, White M, Bennett AF (2018) Identifying wildlife corridors for the restoration of regional habitat connectivity: a multispecies approach and comparison of resistance surfaces. PLoS One 13:e0206071. https://doi.org/10.1371/journal.pone.0206071

Lobo JM, Jiménez-Valverde A, Hortal J (2010) The uncertain nature of absences and their importance in species distribution modelling. Ecography 33:103–114. https://doi.org/10.1111/j.1600-0587.2009.06039.x

Louis EJ, Dempster ER (1987) An Exact Test for Hardy-Weinberg and Multiple Alleles. Biometrics 43:805-811. https://doi.org/10.2307/2531534

Maunder M, Leiva A, Santiago-Valentín E, et al (2008) Plant conservation in the Caribbean Island biodiversity hotspot. Bot Rev 74:197–207

Mod HK, Scherrer D, Luoto M, Guisan A (2016) What we use is not what we know: environmental predictors in plant distribution models. J Veg Sci 27:1308–1322. https://doi.org/10.1111/jvs.12444

Naimi B, Araújo MB (2016) sdm: a reproducible and extensible R platform for species distribution modelling. Ecography 39:368–375. https://doi.org/10.1111/ecog.01881

Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. Ann Hum Genet 47:253–259. https://doi.org/10.1111/j.1469-1809.1983.tb00993.x

Oldfield S, Newton AC (2012) Integrated conservation of tree species by botanic gardens: a reference manual. Botanic Gardens Conservation International, Richmond, UK

Peakall R, Smouse PE (2012) GenALEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. Bioinformatics 28:2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Peakall R, Smouse PE (2006) genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. Annu Rev Ecol Evol Syst 37:187–214. https://doi.org/10.1146/annurev.ecolsys.37.091305.110215

Phillips SJ, Dudík M, Elith J, et al (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecol Appl 19:181–197. https://doi.org/10.1890/07-2153.1

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genet Soc Am 155:946–959. https://doi.org/10.1007/s10681-008-9788-0

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249. https://doi.org/10.1093/oxfordjournals.jhered.a111573

Rivers M, Beech E, Murphy L, Oldfield S (2016) The Red List of Magnoliaceae - revised and extended. Botanic Gardens Conservation International

Rivers MC, Brummitt NA, Nic Lughadha E, Meagher TR (2014) Do species conservation assessments capture genetic diversity? Glob Ecol Conserv 2:81–87. https://doi.org/10.1016/j.gecco.2014.08.005

Roberge J, Angelstam P (2004) Usefulness of the umbrella species concept. Conserv Biol 18:76–85

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 4:137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Rousset F (2008) GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Resour 8:103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Sanchez M, Ingrouille MJ, Cowan RS, et al (2014) Spatial structure and genetic diversity of natural populations of the Caribbean pine, *Pinus caribaea* var. *bahamensis* (Pinaceae), in the Bahaman archipelago. Bot J Linn Soc 174:359–383. https://doi.org/10.1111/boj.12146

Sangermano F, Bol L, Galvis P, et al (2015a) Habitat suitability and protection status of four species of amphibians in the Dominican Republic. Appl Geogr 63:55–65. https://doi.org/10.1016/j.apgeog.2015.06.002

Sangermano F, Bol L, Galvis P, et al (2015b) Forest baseline and deforestation map of the Dominican Republic through the analysis of time series of MODIS data. Data Br 4:363–367. https://doi.org/10.1016/j.dib.2015.06.006

Seddon PJ, Griffiths CJ, Soorae PS, Armstrong DP (2014) Reversing defaunation: restoring species in a changing world. Science (80-) 345:406–412

Slatkin M (2008) Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. Nat Rev Genet 9:477–485

Sorensen FC (1999) Relationship between self-fertility, allocation of growth, and inbreeding depression in three coniferous species. Evolution (N Y) 53:417–425. https://doi.org/10.1111/j.1558-5646.1999.tb03777.x

Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. Proc Natl Acad Sci USA 101:15261–15264. https://doi.org/10.1073/pnas.0403809101

Tamaki I, Setsuko S, Tomaru N (2009) Estimation of outcrossing rates at hierarchical levels of fruits, individuals, populations and species in *Magnolia stellata*. Heredity (Edinb) 102:381–388. https://doi.org/10.1038/hdy.2008.128

Tamaki I, Setsuko S, Tomaru N (2008) Genetic variation and differentiation in populations of a threatened tree, *Magnolia stellata*: Factors influencing the level of within-population genetic variation. Heredity 100:415–423. https://doi.org/10.1038/sj.hdy.6801097

Tamaki I, Setsuko S, Tomaru N (2016) Genetic diversity and structure of remnant *Magnolia stellata* populations affected by anthropogenic pressures and a conservation strategy for maintaining their current genetic diversity. Conserv Genet 17:715–725. https://doi.org/10.1007/s10592-016-0817-6

Tamaki I, Tani S, Setsuko S, et al (2017) Reduced incompatibility in the production of second generation hybrids between two *Magnolia* species revealed by bayesian gene dispersal modeling. Am J Bot 104:1546–1555. https://doi.org/10.3732/ajb.1700138

Thien LB (1974) Floral biology of Magnolia. Am J Bot 61:1037. https://doi.org/10.2307/2441921

Title PO, Bemmels JB (2018) ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. Ecography 41:291–307. https://doi.org/10.1111/ecog.02880

Urbina-Cardona N, Blair ME, Londoño MC, et al (2019) Species distribution modeling in Latin America: a 25-year retrospective review. Trop Conserv Sci 12:1–19. https://doi.org/10.1177/1940082919854058

Väli Ü, Einarsson A, Waits L, Ellegren H (2008) To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? Mol Ecol 17:3808–3817. https://doi.org/10.1111/j.1365-294X.2008.03876.x

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

VanDerWal J, Shoo LP, Graham C, Williams SE (2009) Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? Ecol Modell 220:589–594. https://doi.org/10.1016/j.ecolmodel.2008.11.010

Vartia S, Collins PC, Cross TF, et al (2014) Multiplexing with three-primer PCR for rapid and economical microsatellite validation. Hereditas 151:43–54. https://doi.org/10.1111/hrd2.00044

Vellend M, Geber MA (2005) Connections between species diversity and genetic diversity. Ecol Lett 8:767–781

Veltjen E (2020) The Caribbean *Magnolia* species (Magnoliaceae): assessment of the genetic diversity and the underlying evolutionary history. PhD thesis. Ghent University, Belgium. pp. 382

Veltjen E, Asselman P, Hernández Rodríguez M, et al (2019) Genetic patterns in Neotropical Magnolias (Magnoliaceae) using de novo developed microsatellite markers. Heredity 122:485–500. https://doi.org/10.1038/s41437-018-0151-5

Venter O, Sanderson EW, Magrach A, et al (2018) Last of the Wild Project, version 3 (LWP-3): 2009 human footprint, 2018 release

Venter O, Sanderson EW, Magrach A, et al (2016) Global terrestrial human footprint maps for 1993 and 2009. Sci Data 3:1–10. https://doi.org/10.1038/sdata.2016.67

Villero D, Pla M, Camps D, et al (2017) Integrating species distribution modelling into decision-making to inform conservation actions. Biodivers Conserv 26:251–271

Wahlund S (1928) Zusammensetzung von populationen und korrelationerscheinungen vom standpunkt der vererbungslehre aus betrachtet. Hereditas 11:65–106. https://doi.org/10.1111/j.1601-5223.1928.tb02483.x

Waples RS (2015) Testing for Hardy–Weinberg proportions: have we lost the plot? J Hered 106:1–19. https://doi.org/10.1093/jhered/esu062

Ward SM, Jasieniuk M (2009) Review: sampling weedy and invasive plant populations for genetic diversity analysis . Weed Sci 57:593–602. https://doi.org/10.1614/ws-09-082.1

Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33:607–611. https://doi.org/10.1111/j.1600-0587.2009.06142.x

Weckworth B V., Hebblewhite M, Mariani S, Musiani M (2018) Lines on a map: conservation units, meta-population dynamics, and recovery of woodland caribou in Canada. Ecosphere 9:e02323. https://doi.org/10.1002/ecs2.2323

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

Wheeler L (2015) *Magnolia domingensis*. In: IUCN Red List Threat Species 2015 eT193937A2291464. https://dx.doi.org/10.2305/IUCN.UK.2015-2.RLTS.T193937A2291464.en. Accessed 20 May 2021

Wheeler L, Rivers MC (2020) *Magnolia ekmanii*. In: he IUCN Red List Threat Species 2020 eT193938A121363857. https://dx.doi.org/10.2305/IUCN.UK.2020-3.RLTS.T193938A121363857.fr. Accessed 20 May 2021

Whiteley AR, Fitzpatrick SW, Funk WC, Tallmon DA (2014) Genetic rescue to the rescue. Trends Ecol Evol. https://doi.org/10.1016/j.tree.2014.10.009

Williams JN, Seo C, Thorne J, et al (2009) Using species distribution models to predict new occurrences for rare plants. Divers Distrib 15:565–576. https://doi.org/10.1111/j.1472-4642.2009.00567.x

Wunderlich RF, Lin Y-P, Anthony J, Petway JR (2019) Two alternative evaluation metrics to replace the true skill statistic in the assessment of species distribution models. Nat Conserv 35:97–116. https://doi.org/10.3897/natureconservation.35.33918

Article title: An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola.

Author names: Tim Claerhout, Emily Veltjen, Diederik Strubbe, Ramón Elías Castillo Torres, Marie-Stéphanie Samain, Lars Chatrou
Corresponding author, affiliation and e-mail address: Tim Claerhout; Systematics and Evolutionary Botany lab, Research Group Spermatophytes, Department of Biology; Tim.Claerhout@UGent.be

---

Online Resource 1 Herbarium material employed as occurrence data for species distribution modelling analyses with their respective latitudinal and longitudinal positions as decimal degrees. Herbarium acronyms follow Thiers ([continuously updated]).

Thiers B, (continuously updated). Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. http://sweetgum.nybg.org/science/ih/

SPECIES	**ACRONYM**	**COLLECTOR(S)**	**HERBARIUM N°**	**DATE OF COLLECTION**	**LATITUDE**	**LONGITUDE**
Magnolia domingensis Urb.	JBSD	Veloz A. & Castillo R. E.	3117	9/05/2004	18.43361111	-70.3075
Magnolia ekmanii Urb.	FLAS	J. C. Timyan	39	7/06/2013	18.37339	-74.29497
Magnolia hamorii Howard	US	Acevedo-Rdgz. P. et al.	13814	20/05/2004	18.10755556	-71.19186111
Magnolia hamorii Howard	JBSD	Clase T. et al.	8416	1/03/2014	18.11980556	-71.22669444
Magnolia hamorii Howard	JBSD	Clase T. et al.	8034	30/07/2013	18.110115	-71.220421
Magnolia hamorii Howard	JBSD	Clase T. et al.	3541	4/04/2003	18.067638	-71.182667
Magnolia hamorii Howard	JBSD	Judd W. S. et al.	8101	31/05/2006	18.11072222	-71.22691667
Magnolia hamorii Howard	JBSD	Skean J. D. et al.	4304	31/05/2006	18.11497222	-71.23155556
Magnolia pallescens Urb. & Ekman	JBSD	Clase T. & Encarnacion W.	6767	3/03/2011	18.836189	-70.682663
Magnolia pallescens Urb. & Ekman	JBSD	Clase T. et al.	6415	23/09/2010	18.8405	-70.72572222
Magnolia pallescens Urb. & Ekman	JBSD	Peguero B. et al.	727	10/06/1998	18.836517	-70.535853
Magnolia pallescens Urb. & Ekman	JBSD	Peguero B. et al.	735	11/06/1998	18.839738	-70.527617

Article title: An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola.
Author names: Tim Claerhout, Emily Veltjen, Diederik Strubbe, Ramón Elías Castillo Torres, Marie-Stéphanie Samain, Lars Chatrou
Corresponding author, affiliation and e-mail address: Tim Claerhout; Systematics and Evolutionary Botany lab, Research Group Spermatophytes, Department of Biology; Tim.Claerhout@UGent.be

Online Resource 2 Sample information regarding the 24 test individuals used for SSR marker testing. **Ref. N°**: reference number to uniquely identify an individual. **OD (260/230)**: ratio of absorbance (i.e. Optical Density ratio) at 260 and 230 nm as a proxy for DNA purity. A ratio between 2.0 and 2.2 is considered pure. **OD (260/280)**: ratio of absorbance at 260 and 280 nm as a proxy for DNA purity. A ratio of ~1.8 is considered pure. **ID**: three letter code to uniquely identify a species. **Loc. ID**: three letter code to uniquely identify locality of origin.

REF. N°	OD (260/230)	OD (260/280)	ID	LOC. ID
MA354	2.02	2.09	PAL	CAS
MA437	1.96	2.04	PAL	CAS
MA468	1.99	1.98	PAL	SAL
MA499	2.02	2.13	PAL	SAL
MA716	2.05	1.76	PAL	ENT
MA743	2.09	1.97	PAL	ENT
MA763	2.02	1.86	PAL	MON
MA780	2.05	1.99	PAL	MON
MA795	2.05	1.77	HAM	COR
MA809	1.97	1.71	HAM	COR
MA822	2.05	1.89	HAM	COR
MA839	1.97	2.04	HAM	COR
MA849	2.07	1.88	HAM	CAC
MA862	2.09	1.99	HAM	CAC
MA869	1.99	2.01	HAM	CAC
MA882	2.06	1.93	HAM	CAC
MA891	2.06	1.91	DOM	BAR
MA903	2.03	1.91	DOM	BAR
MA904	2.01	1.88	DOM	BAR
MA906	2.05	1.93	DOM	BAR
MA908	2.05	1.95	DOM	ROD
MA922	2.04	1.99	DOM	ROD
MA931	2.04	1.96	DOM	ROD
MA949	2.05	1.94	DOM	ROD

Article title: An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola.
Author names: Tim Claerhout, Emily Veltjen, Diederik Strubbe, Ramón Elías Castillo Torres, Marie-Stéphanie Samain, Lars Chatrou
Corresponding author, affiliation and e-mail address: Tim Claerhout; Systematics and Evolutionary Botany lab, Research Group Spermatophytes, Department of Biology; Tim.Claerhout@UGent.be

---

Online Resource 3 Total set of 25 microsatellite markers that were evaluated for their performance and for which primer information is given. This resulted in a final set of 16 markers. Primers were developed on *Magnolia lacondonica* (MA39), *M. mayae* (MA40), *M. dealbata* (MA41) and *M. cubensis* subsp. *acunae* (MA42). Markers indicated with an asterisk (*) were removed due to non-random allelic associations (Linkage Disequilibrium; LD). Markers were obtained from Veltjen et al. (2019).

LOCUS NAME	PRIMER SEQUENCE (5'-3')	DYE	REPEAT MOTIF	GENBANK ACCESSION NUMBER	FINAL MARKER
MA39_023	F: ATCACGCATCTGCACAGACA R: GGACAACGAACGTCTGGCTA	FAM	(AG)7	MH923371	☒
MA39_165	F: AATGTAGTGGGTCCGGCTTC R: CCAAACCATGTGCGTCCTTG	PET	(TC)18	MH923375	☒
MA39_185	F: CGGGTGTTGTAGATGACGCT R: AAGACACGGAATGGGACGAG	PET	(AG)15	MH923377	☒
MA39_199	F: CGCCCACATCTACCTCTTCG R: TCCAGGAGTTTCTGTGCACC	FAM	(GGA)5	MH923378	☐
MA39_259	F: TGATAGAGTGGGATGGCGGA R: TGCTGCTTTGAGGCCTGTTA	VIC	(CT)11	MH923380	☐
MA40_045*	F: TTGTGGGCCAAGCTCGATAG R: ATTGTGGCATGTACCTCGCA	VIC	(TC)13	MH923387	☐
MA40_223	F: TTCAGTGGCTGGAGCTTCAG R: GGAGCATCTTGGCCTTTGGA	VIC	(GAT)5	MH923391	☐
MA40_282	F: TCTCTTTCCCTCCGTCCTCC R: TCTTCCGGCTTCATGTCGTC	FAM	(GA)15	MH923392	☒
MA41_373	F: GCGCCCAATCAGAACACAAC R: GGGAAGAGCTTCTTTCGCCA	NED	(CT)16	MH923396	☒
MA42_001	F: ATCCGACCCAACATGGTGAC R: AGCCGAGTCTGAGCTGAGTA	PET	(TC)11	MH923397	☒

MA42_059	F: AGGGACTCGGCATCTATGGA R: GAGTCGACTCAGCAACTCCC	PET	(AG)8	MH923399	☒
MA42_077	F: GAGACATGGAACCCACACGT R: CTGGTGGTCTAGCCGATCTG	FAM	(AG)8	MH923402	☐
MA42_083	F: GTCTTCCACGGGAGCAAGAG R: CGAGTTGGACCCAGTGAGTC	VIC	(GAA)17	MH923403	☐
MA42_126	F: CACATCGTCCGTCCAGACAT R: TCGCCTAGCCAATAGTCTGC	FAM	(AT)9	MH923406	☒
MA42_203	F: TGAAGAACACAGGCCATGGA R: GAGAGGTGCTTCACGGGTAG	FAM	(TC)16	MH923412	☒
MA42_231	F: GGGTGCGAAATGTGCATCAA R: GGGCCAGTGAGCATTAGAGC	VIC	(AG)14	MH923413	☒
MA42_241	F: GGGTACCCTATGGTCCAACC R: GTCCGACTAAGGCCCATTGT	NED	(CA)11	MH923414	☒
MA42_255	F: ACGTGGGTCGAGGATCAAGT R: GGACCCACCTCCAACAGATC	PET	(AG)14	MH923417	☒
MA42_293	F: TGCAACTGAGACGAGTTGGG R: GGTACGGACTAGGGTACAGGT	NED	(GA)16	MH923421	☒
MA42_397	F: TAGTAGCAGGGTCCCTCCTC R: TCCATTCATTAGGGTGGGCA	NED	(TC)20	MH923426	☒
MA42_421	F: GACAGCAGACCTGACCGATT R: GACCAGTGCATCCCATCAAA	VIC	(TC)10	MH923428	☒
MA42_471	F: TGATGAAGAGCCCAGATCGTC R: TGGCCTTGTTCTCCATACGT	FAM, VIC	(GA)16	MH923429	☐
MA42_472	F: AGAGTTACACATGCAAACCCG R: TGATGTTGTTGCTCGGCTGA	PET	(AG)17	MH923430	☒
MA42_481	F: CGATCTGAGTCCGCAAGAGT R:GACGCAGAAATCTCAGCAAGA	NED	(TC)15	MH923431	☐
MA42_495	F: TGCATCTCCTCATCCTCCCA R: ACGCCATTCAATTACCTACGG	PET	(GA)26	MH923433	☐

Article title: An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola.

Author names: Tim Claerhout, Emily Veltjen, Diederik Strubbe, Ramón Elías Castillo Torres, Marie-Stéphanie Samain, Lars Chatrou

Corresponding author, affiliation and e-mail address: Tim Claerhout; Systematics and Evolutionary Botany lab, Research Group Spermatophytes, Department of Biology; Tim.Claerhout@UGent.be

---

Online Resource 4 Final set of predictor variables included in the species distribution modelling analyses of the genus *Magnolia* in Hispaniola. Sources: **A:** Karger et al. (2017); **B:** Title and Bemmels (2018); **C:** Venter et al. (2016, 2018).

N°	VARIABLE	DESCRIPTION	UNIT	SOURCE
1	Bio1	Annual mean temperature.	°C*10	A
2	Bio2	Mean diurnal range.	°C	A
3	Bio3	Isothermality.	-	A
4	Bio4	Temperature seasonality.	SD	A
5	Bio8	Mean temperature of wettest quarter.	°C*10	A
6	Bio12	Annual precipitation.	mm/year	A
7	Bio15	Precipitation seasonality.	Coefficient of variation	A
8	Bio16	Precipitation of wettest quarter.	mm/quarter	A
9	Bio18	Precipitation of warmest quarter.	mm/quarter	A
10	AnnualPET	Annual potential evapotranspiration: a measure of the ability of the atmosphere to remove water through evapotranspiration processes, given unlimited moisture.	mm/year	B
11	AridityIndexT	Thornthwaite aridity index: Index of the degree of water deficit below water need.	-	B
12	ClimaticMoisture Index	A metric of relative wetness and aridity.	-	B
13	Continentality	Average temp. of warmest month – average temp. of coldest month.	°C	B
14	EmbergerQ	Emberger's pluviothermic quotient: a metric that was designed to differentiate among Mediterranean type climates.	-	B
15	MonthCountBy Temp10	Count of the number of months with mean temp. greater than 10°C.	months	B
16	PETColdestQuarter	Mean monthly PET of coldest quarter.	mm/month	B
17	PETDriestQuarter	Mean monthly PET of driest quarter.	mm/month	B
18	PETWettestQuarter	Mean monthly PET of wettest quarter.	mm/month	B
19	Tri	Terrain roughness index.	-	B
20	TopoWet	SAGA-GIS topographic wetness index.	-	B
21	Footprint	Quantitative measure of human alteration of terrestrial environments in 2009, based on human population size, land use and infrastructure.	-	C

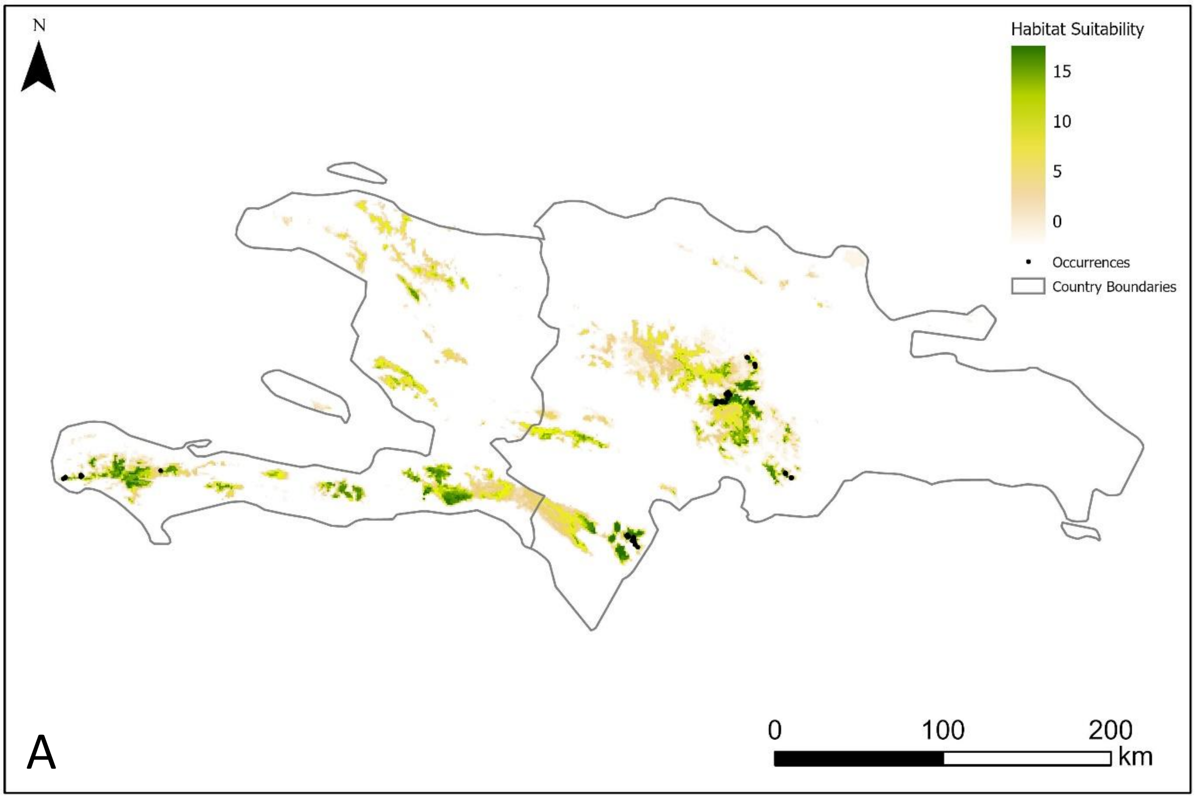
Article title: An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola.

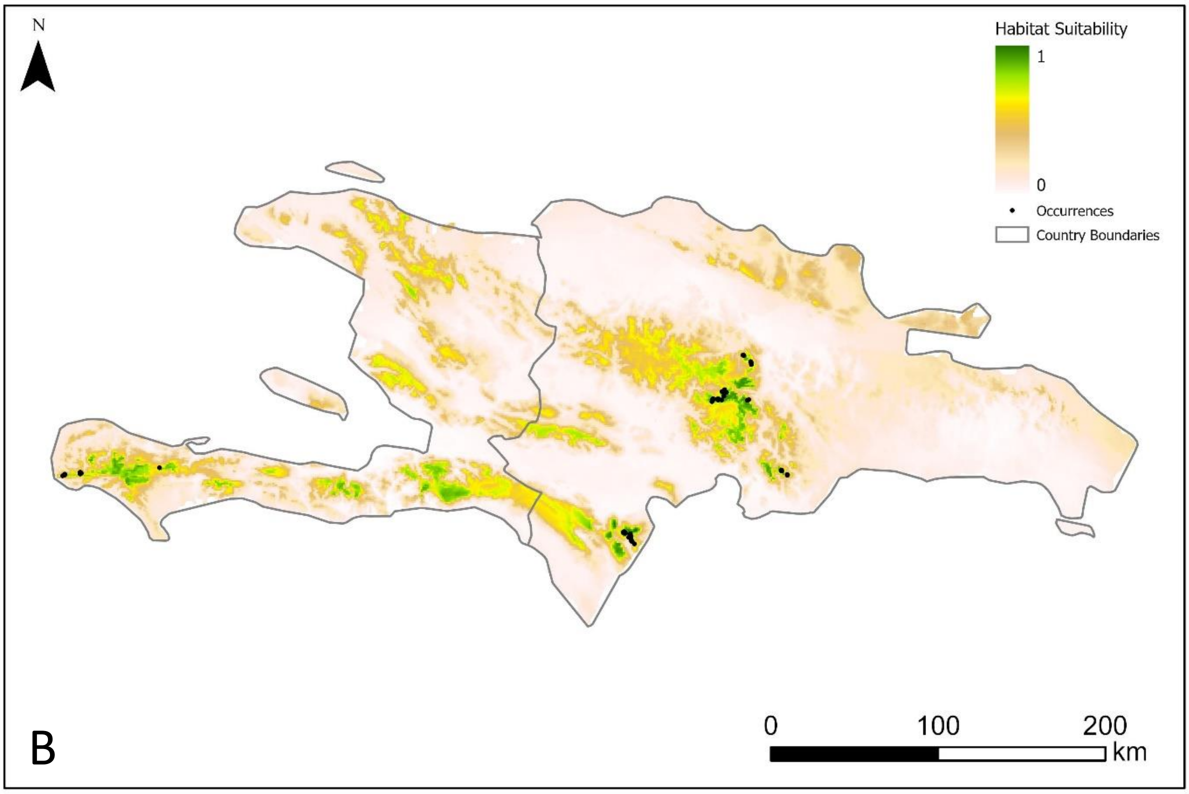
Author names: Tim Claerhout, Emily Veltjen, Diederik Strubbe, Ramón Elías Castillo Torres, Marie-Stéphanie Samain, Lars Chatrou

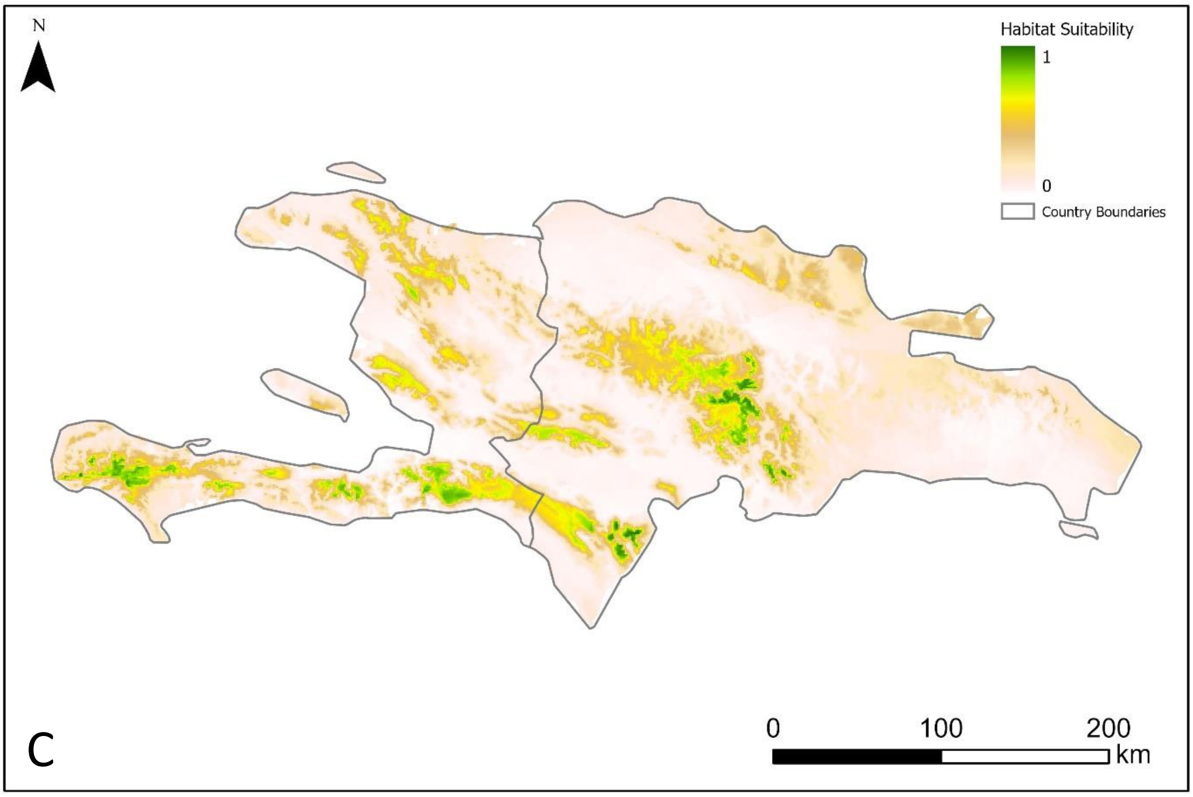
Corresponding author, affiliation and e-mail address: Tim Claerhout; Systematics and Evolutionary Botany lab, Research Group Spermatophytes, Department of Biology; Tim.Claerhout@UGent.be

---

Online Resource 5 Maps visualizing the potential distribution and habitat suitability for the Magnolias of Hispaniola. **A**. Ensemble-based distribution map for using "vote counting." Habitat suitability ranges from zero to fifteen, indicating the number of ensemble analyses that labelled a pixel as suitable (i.e. TSS > 0.7). **B**. Ensemble-based distribution map for using "unweighted averaging." Habitat suitability ranges from zero to one, indicating the average TSS-value of a pixel. **C**. The same map as **B**, but without plotted occurrences.




N
Habitat Suitability
1
0
Occurrences
Country Boundaries
0
100
200
km
B


N
Habitat Suitability
1
0
Country Boundaries
0
100
200
km
C

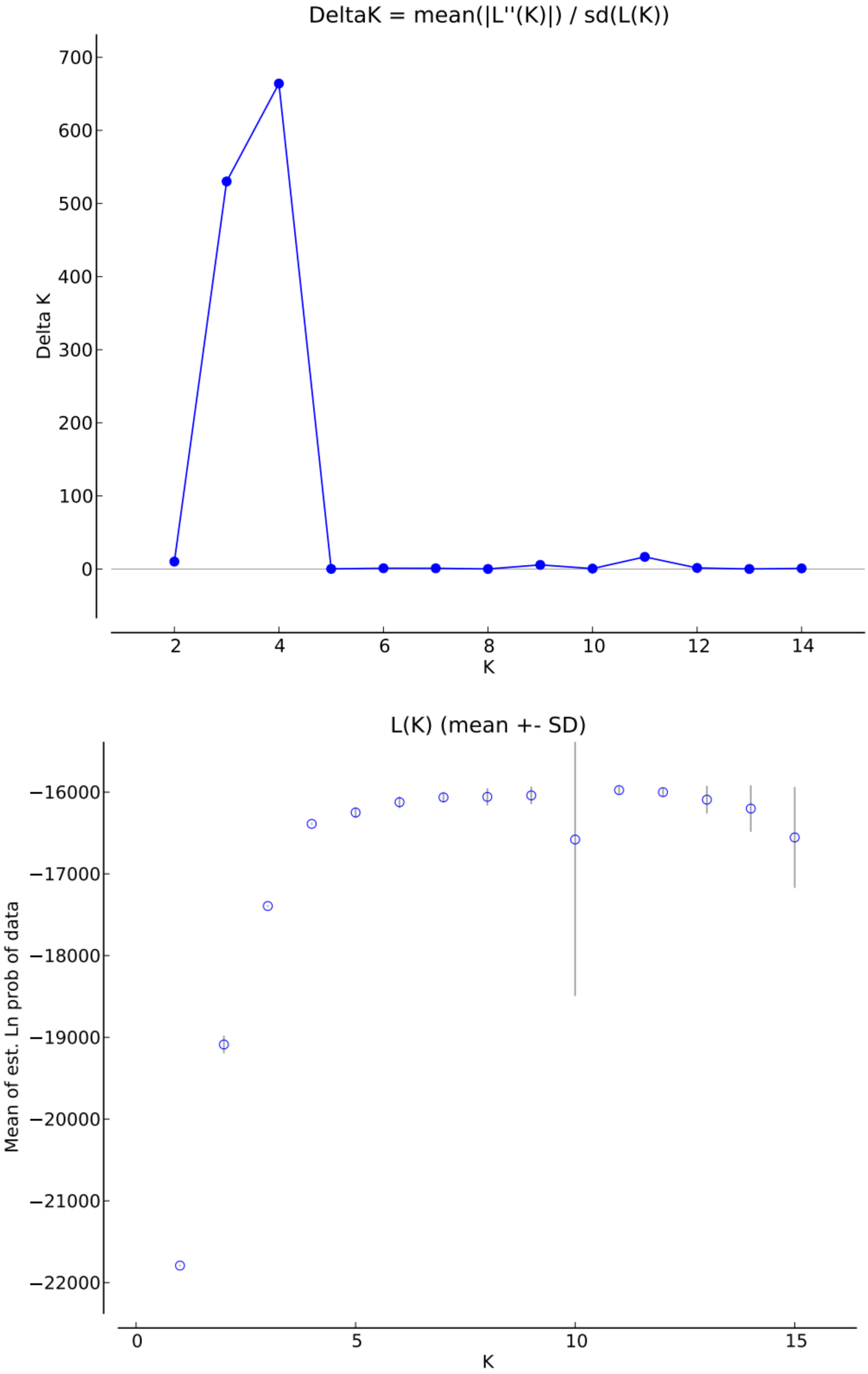
Article title: An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola.

Author names: Tim Claerhout, Emily Veltjen, Diederik Strubbe, Ramón Elías Castillo Torres, Marie-Stéphanie Samain, Lars Chatrou

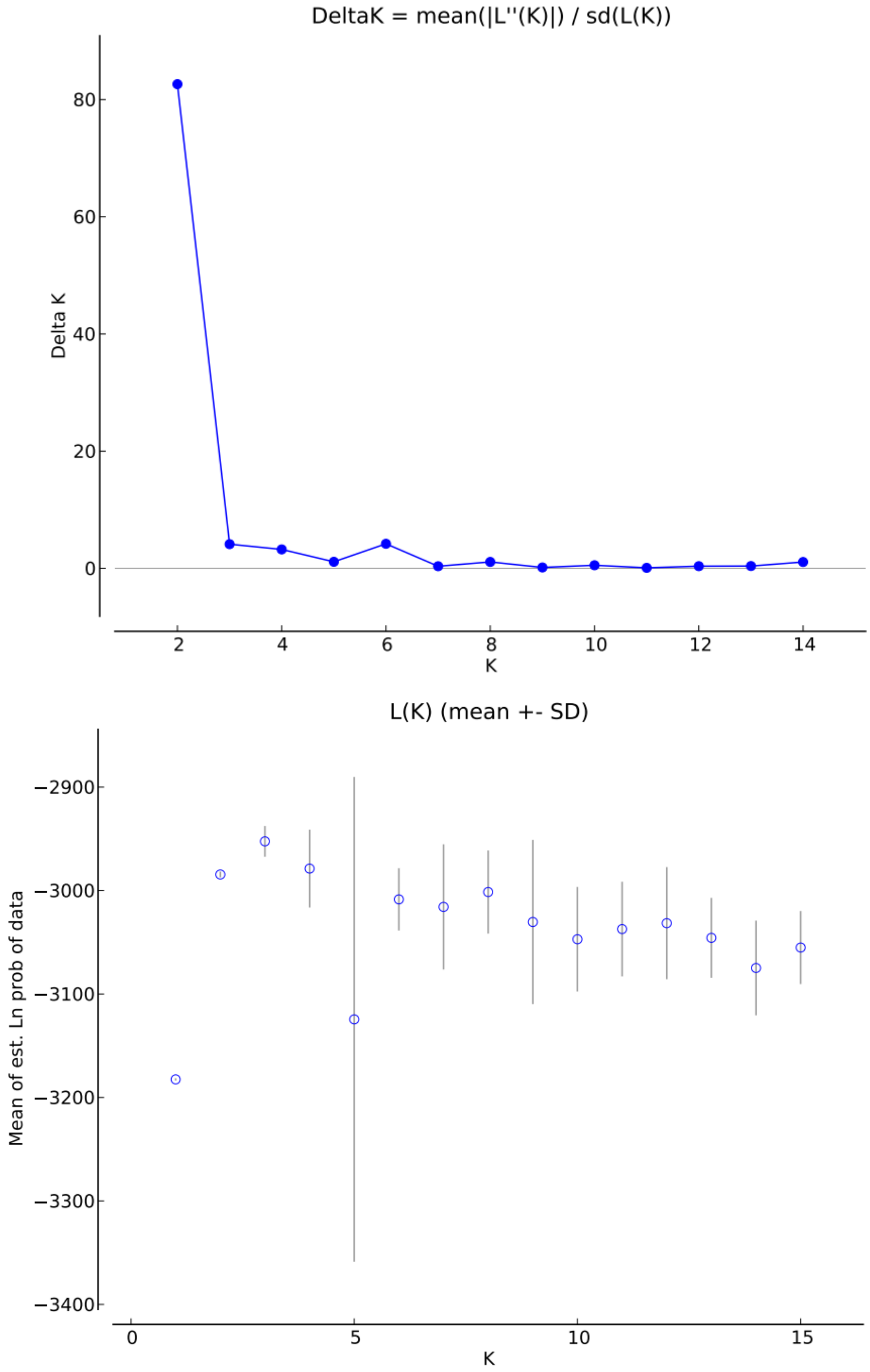
Corresponding author, affiliation and e-mail address: Tim Claerhout; Systematics and Evolutionary Botany lab, Research Group Spermatophytes, Department of Biology; Tim.Claerhout@UGent.be

Online Resource 6 STRUCTURE ΔK (Evanno et al. 2005) and mean LnK plots for all four analyses. Individuals are genotyped for 16 SSR microsatellite markers. **A**. Dataset comprising all 417 individuals of every species. **B**. *Magnolia domingensis*. **C**. *Magnolia hamorii.* **D**. *Magnolia pallescens*.

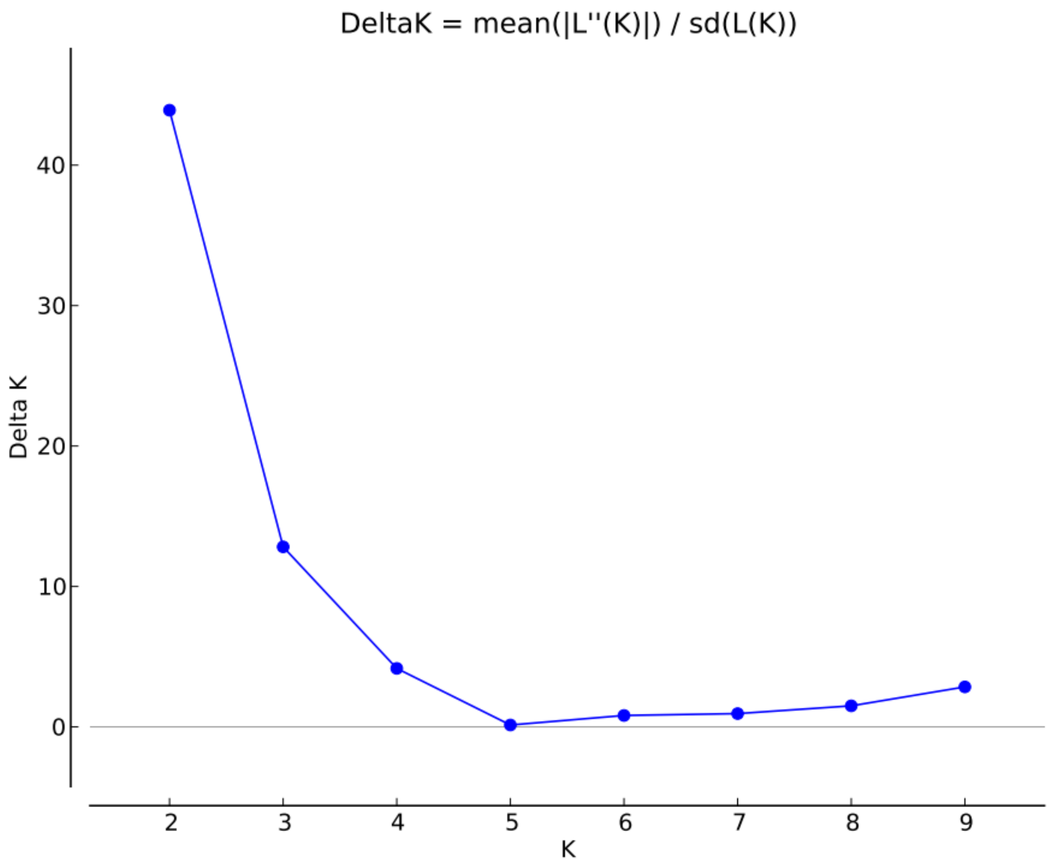
A. *Magnolia spp.* from the Dominican Republic.

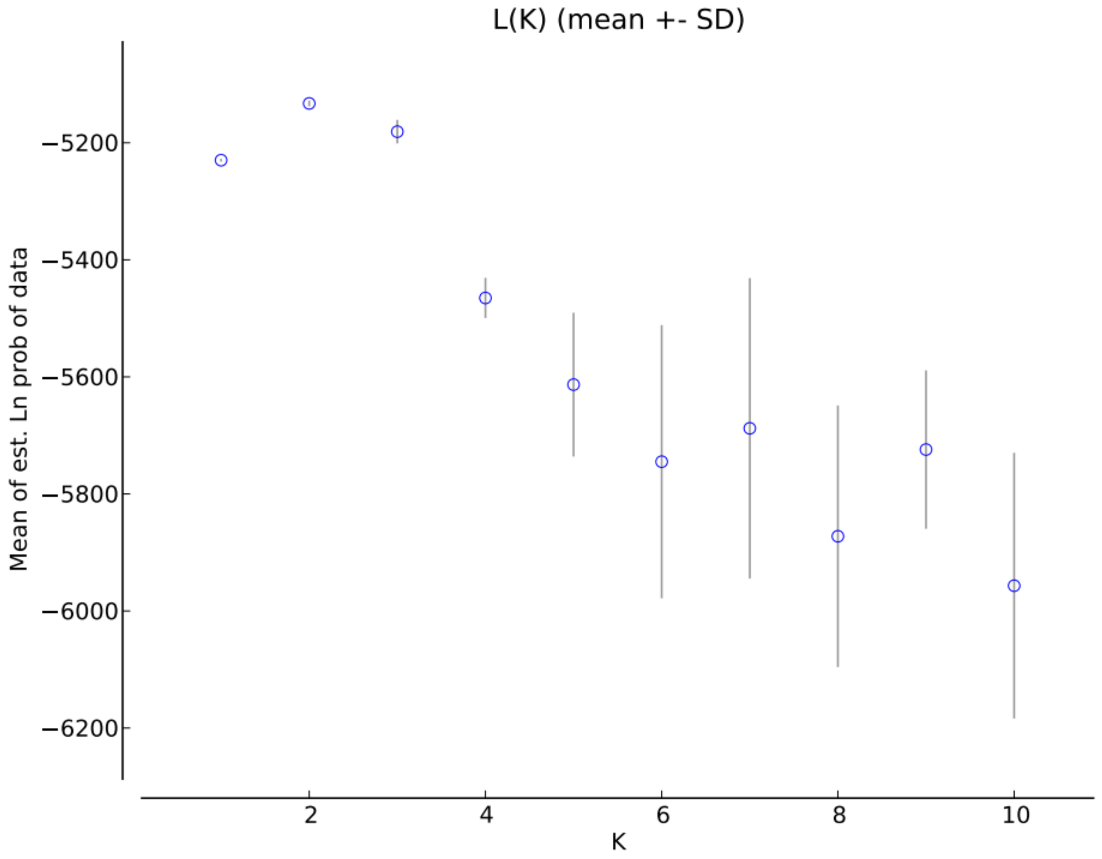


B. ***Magnolia domingensis.***

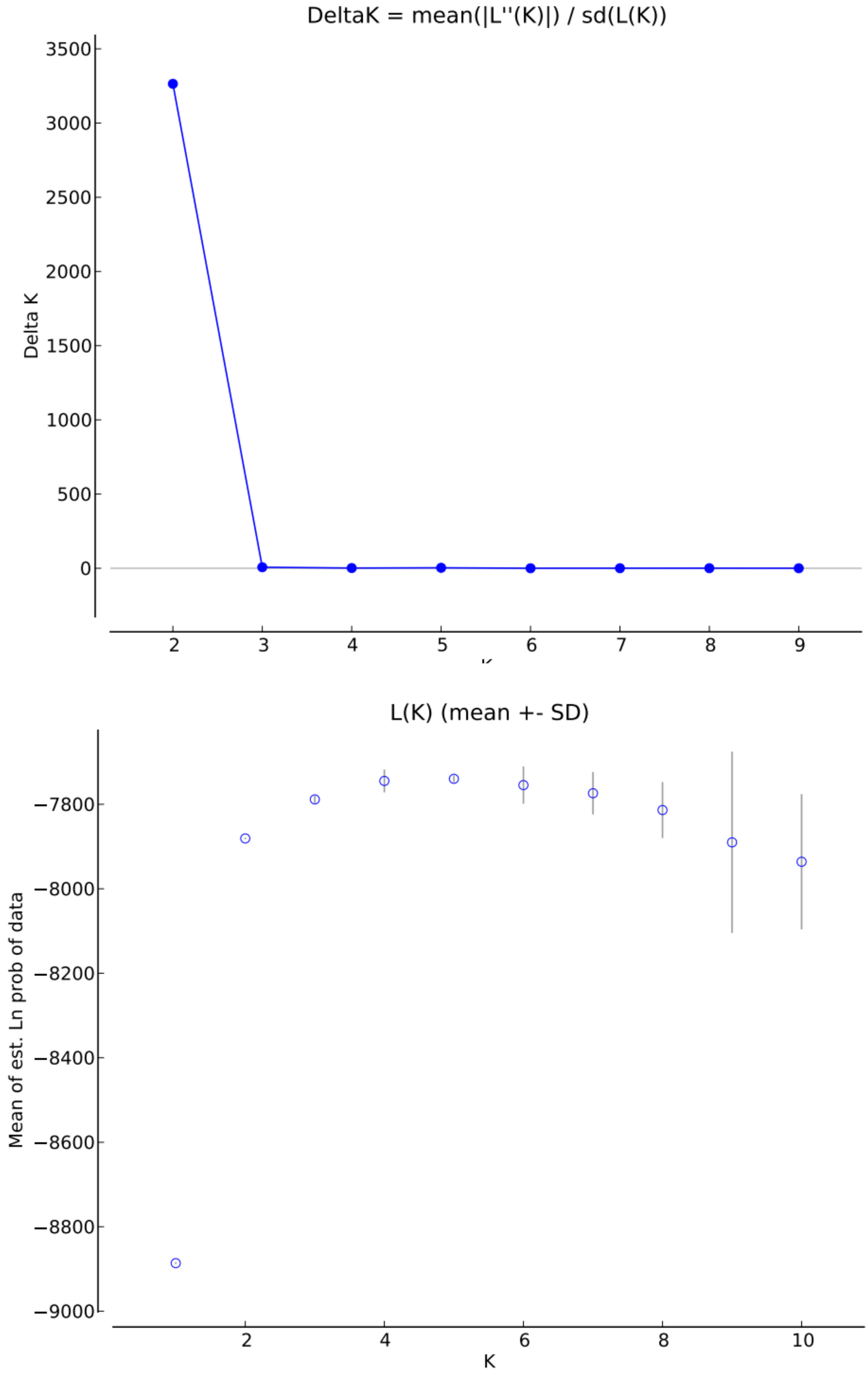


C. *Magnolia hamorii.*





D. ***Magnolia pallescens.***



Article title: An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola.

Author names: Tim Claerhout, Emily Veltjen, Diederik Strubbe, Ramón Elías Castillo Torres, Marie-Stéphanie Samain, Lars Chatrou

Corresponding author, affiliation and e-mail address: Tim Claerhout; Systematics and Evolutionary Botany lab, Research Group Spermatophytes, Department of Biology; Tim.Claerhout@UGent.be

---

Online Resource 7 DAPC analyses results. Tables show the distribution of individuals across X genetic clusters. Between brackets: number of Principal Component Axes retained, based on the lowest mean squared error. **LOC/POP**: three letter code to uniquely identify locality or population, following Table 2.

A. Full dataset (60)

LOC | CLUSTERS	1	2	3	4
BAR	0	0	0	36
ROD	0	0	0	50
CAC	0	52	0	0
COR	0	52	0	0
LAG	0	7	0	0
TON	0	7	0	0
CAS	0	0	40	0
SAL	0	0	40	0
ENT	61	0	0	0
MON	41	0	0	0
GUA	15	0	0	0
RAN	16	0	0	0

B. *Magnolia domingensis* (10)

LOC | CLUSTERS	1	2
BAR	34	2
ROD	6	44

C. *Magnolia hamorii* (10)

LOC | CLUSTERS	1	2
CAC	46	6
COR	18	34
LAG	5	2
TON	2	5

D. *Magnolia pallescens* (10)

POP | CLUSTERS	1	2
EBV	80	0
VAL	0	133

E. Ébano Verde (10)

LOC | CLUSTERS	1	2
CAS	37	3
SAL	8	32

F. Valle Nuevo (10)

LOC | CLUSTERS	1	2
ENT	39	22
MON	14	27
GUA	12	3
RAN	1	15

Article title: An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola.
Author names: Tim Claerhout, Emily Veltjen, Diederik Strubbe, Ramón Elías Castillo Torres, Marie-Stéphanie Samain, Lars Chatrou
Corresponding author, affiliation and e-mail address: Tim Claerhout; Systematics and Evolutionary Botany lab, Research Group Spermatophytes, Department of Biology; Tim.Claerhout@UGent.be

---

Online Resource 8 Pairwise fixation and allelic differentiation statistics (FST, GST, G'ST & DJ) for every locality combination. Locality names follow Table 1.

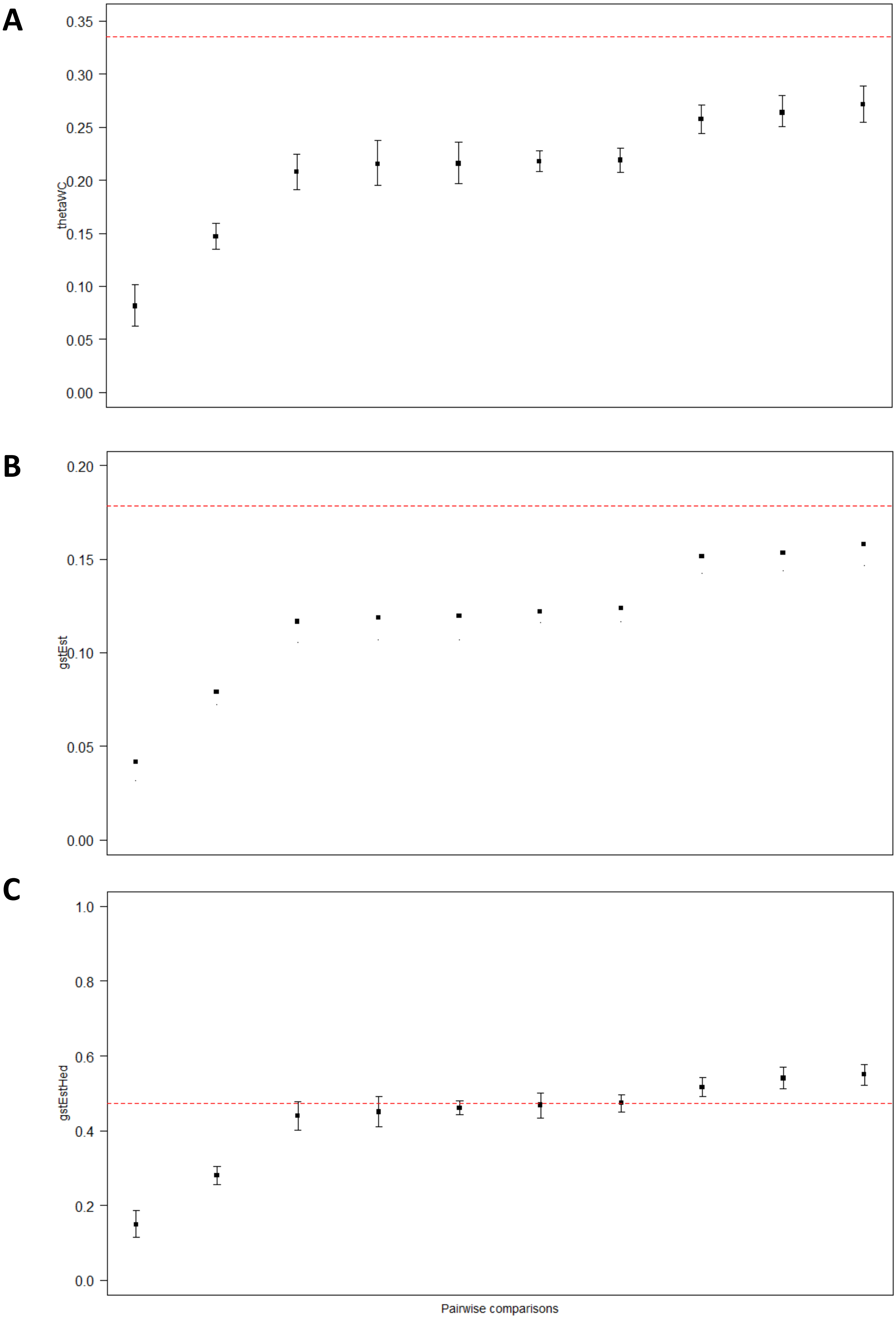
Statistic	Locality	BAR	ROD	CAC	COR	LAG	TON	CAS	SAL	ENT	MON	GUA	RAN
FST	ROD	0.081											
GST		0.042											
G'ST		0.149											
DJ		0.081											
FST	CAC	0.221	0.263										
GST		0.124	0.152										
G'ST		0.477	0.529										
DJ		0.364	0.335										
FST	COR	0.225	0.291	0.050									
GST		0.127	0.171	0.026									
G'ST		0.500	0.609	0.099									
DJ		0.392	0.435	0.069									
FST	LAG	0.212	0.261	0.023	0.069								
GST		0.118	0.144	0.012	0.037								
G'ST		0.457	0.504	0.045	0.141								
DJ		0.328	0.297	0.011	0.074								
FST	TON	0.204	0.272	0.054	0.033	0.047							
GST		0.111	0.149	0.027	0.017	0.024							
G'ST		0.453	0.546	0.107	0.070	0.095							
DJ		0.318	0.358	0.052	0.035	0.015							

FST	CAS	0.216	0.266	0.238	0.232	0.234	0.226					
GST		0.120	0.153	0.136	0.132	0.130	0.122					
G'ST		0.441	0.511	0.487	0.483	0.466	0.461					
DJ		0.257	0.290	0.360	0.369	0.343	0.358					
FST	SAL	0.232	0.299	0.256	0.242	0.256	0.234	0.049				
GST		0.131	0.175	0.147	0.139	0.145	0.128	0.025				
G'ST		0.485	0.589	0.535	0.514	0.526	0.489	0.087				
DJ		0.316	0.393	0.406	0.400	0.416	0.382	0.041				
FST	ENT	0.234	0.289	0.246	0.252	0.262	0.253	0.170	0.143			
GST		0.131	0.169	0.140	0.144	0.146	0.137	0.093	0.077			
G'ST		0.469	0.551	0.491	0.516	0.513	0.508	0.310	0.260			
DJ		0.327	0.344	0.366	0.392	0.367	0.365	0.168	0.128			
FST	MON	0.213	0.264	0.225	0.233	0.237	0.230	0.179	0.154	0.035		
GST		0.119	0.152	0.127	0.133	0.132	0.125	0.098	0.083	0.018		
G'ST		0.432	0.504	0.454	0.484	0.469	0.469	0.334	0.287	0.060		
DJ		0.238	0.261	0.345	0.361	0.341	0.349	0.186	0.145	0.023		
FST	GUA	0.244	0.312	0.259	0.263	0.283	0.265	0.211	0.178	0.039	0.067	
GST		0.143	0.187	0.154	0.158	0.159	0.147	0.120	0.100	0.020	0.036	
G'ST		0.491	0.584	0.517	0.541	0.534	0.518	0.384	0.324	0.064	0.114	
DJ		0.294	0.344	0.360	0.404	0.358	0.381	0.219	0.168	0.024	0.048	
FST	RAN	0.195	0.229	0.202	0.227	0.211	0.211	0.213	0.196	0.055	0.034	0.078
GST		0.108	0.127	0.113	0.129	0.117	0.116	0.118	0.108	0.028	0.017	0.041
G'ST		0.413	0.437	0.420	0.492	0.436	0.458	0.419	0.388	0.097	0.061	0.135
DJ		0.250	0.222	0.321	0.388	0.295	0.339	0.280	0.235	0.031	0.020	0.047

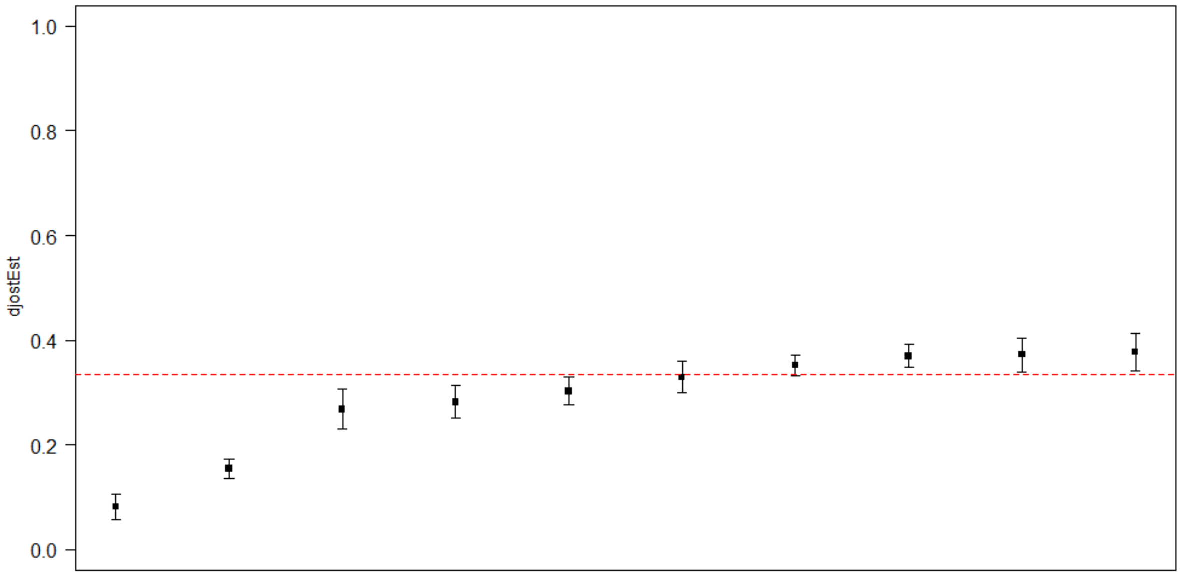
Article title: An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola.

Author names: Tim Claerhout, Emily Veltjen, Diederik Strubbe, Ramón Elías Castillo Torres, Marie-Stéphanie Samain, Lars Chatrou
Corresponding author, affiliation and e-mail address: Tim Claerhout; Systematics and Evolutionary Botany lab, Research Group Spermatophytes, Department of Biology; Tim.Claerhout@UGent.be

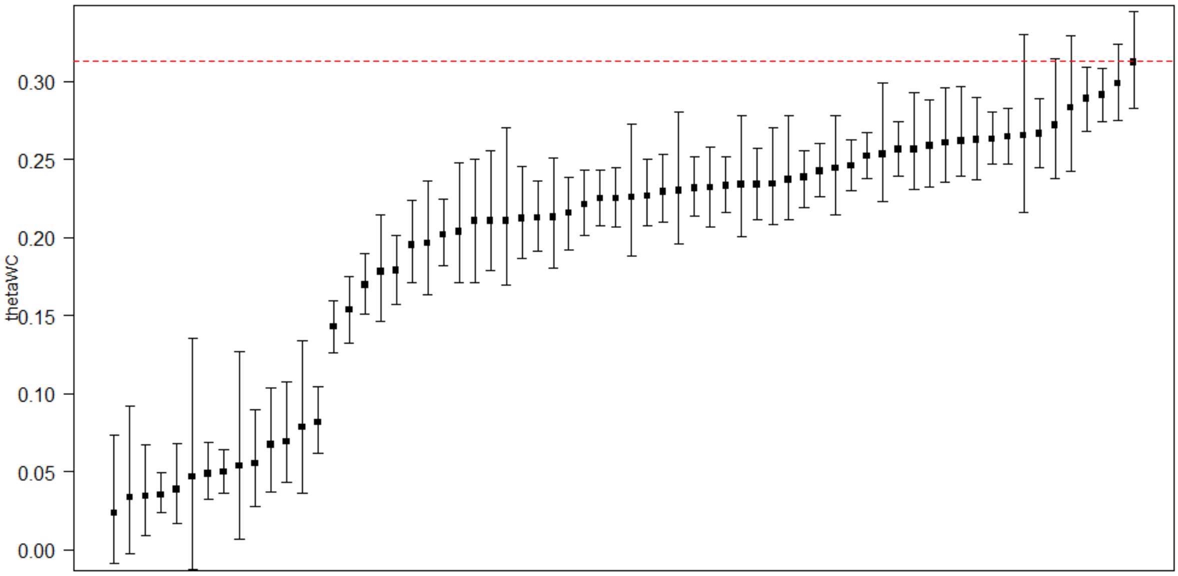
Online Resource 9 95% Confidence Intervals (CIs) of the pairwise fixation and allelic differentiation indices between populations (**A** – **D**) and between localities (**E** – **H**) from Table 4. Red dotted lines represent the global statistic value. **A**. Pairwise population F_{ST} (Weir and Cockerham 1984). **B**. Pairwise population G_{ST} (Nei and Chesser 1983). **C**. Pairwise population G'_{ST} (Hedrick 2005). **D**. Pairwise population Jost's D (D_J; Jost 2008). **E** – **H**: same indices as **A** – **D**, respectively, but between localities.


A
thetaWC
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
B
gstEst
0.00
0.05
0.10
0.15
0.20
C
gstEstHed
0.0
0.2
0.4
0.6
0.8
1.0
Pairwise comparisons

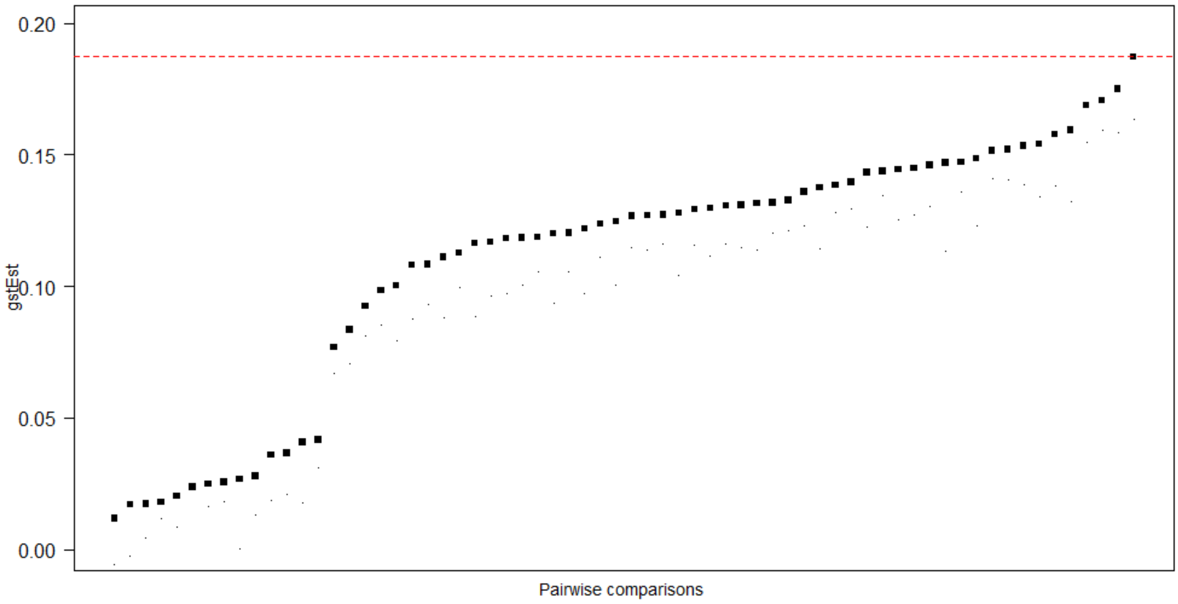
D


djostEst
1.0
0.8
0.6
0.4
0.2
0.0

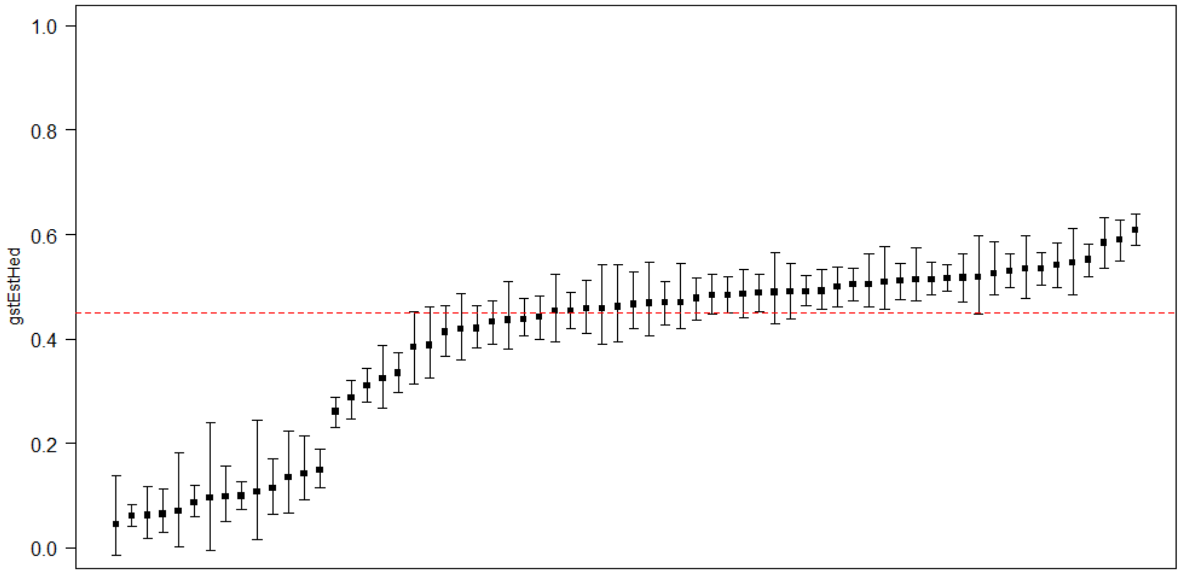
E


thetaWC
0.30
0.25
0.20
0.15
0.10
0.05
0.00

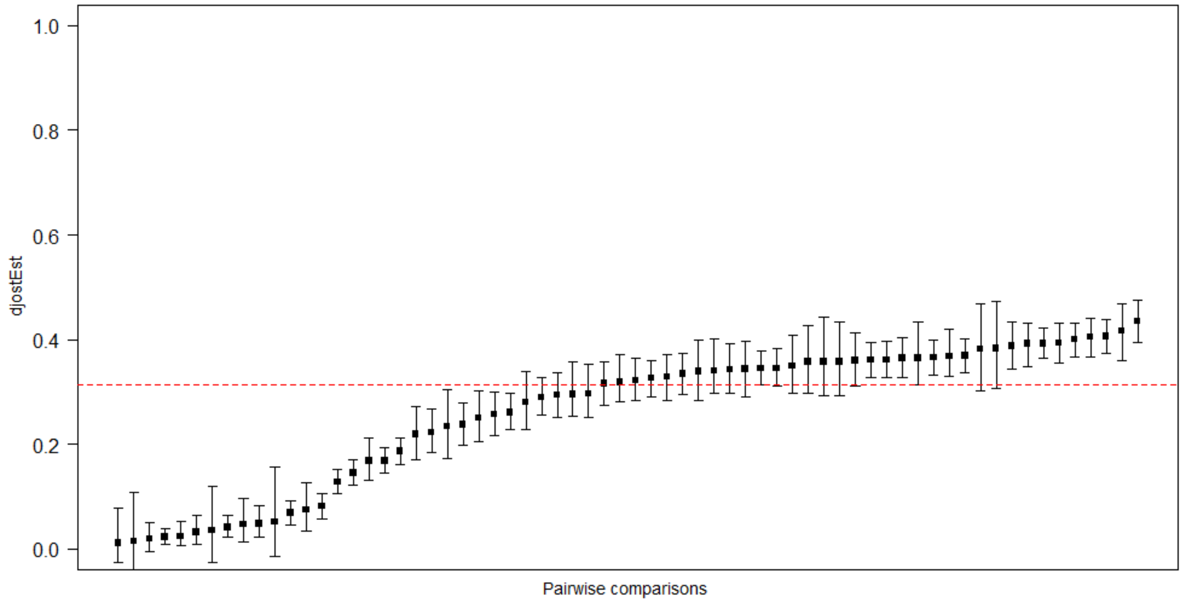
F


gstEst
0.20
0.15
0.10
0.05
0.00
Pairwise comparisons

G


gstEstHed
1.0
0.8
0.6
0.4
0.2
0.0

H


djostEst
1.0
0.8
0.6
0.4
0.2
0.0
Pairwise comparisons

Article title: An integrative study of species distribution modelling and conservation genetics: *Magnolia* in Hispaniola.

Author names: Tim Claerhout, Emily Veltjen, Diederik Strubbe, Ramón Elías Castillo Torres, Marie-Stéphanie Samain, Lars Chatrou

Corresponding author, affiliation and e-mail address: Tim Claerhout; Systematics and Evolutionary Botany lab, Research Group Spermatophytes, Department of Biology; Tim.Claerhout@UGent.be

---

Online Resource 10 Population and locality statistics for the Magnolias of the Dominican Republic across every SSR microsatellite marker. **Population/Locality**. Species, population/locality name and three letter code to uniquely identify a species, population or locality, following Table 1. **N_0**. Number of sampled individuals. **A**. Average allelic diversity. **A_{R-X}**. Allelic richness based on a minimum sample size of 6, 14 and 36 individuals. **A_P**. Number of private alleles between all localities (**A_{PA}**) or between populations only (**A_{PP}**). **H_O**. Average observed heterozygosity. **H_E**. Average expected heterozygosity. **F_{IS}**. Inbreeding coefficient. Significant ($p = 0.05$) deviations from Hardy-Weinberg proportions are indicated with an asterix (*). Monomorphic loci are indicated by “M.” **P**. Number of polymorphic loci. **Pop**: Population statistics for each population and locality as a whole. **SE**: Standard error.

Population	SSR marker	Summary statistics						
		N	A	$A_{R\text{-}36}$	A_{PP} (A_{PA})	H_O	H_E	F_{IS}
Magnolia domingensis Loma Barbacoa (BAR) N_0 = 36 P = 16	MA39_023	36	4	4.000	1 (1)	0.583	0.568	-0.014
	MA39_165	36	2	2.000	0 (0)	0.028	0.027	0.000
	MA39_185	36	5	5.000	0 (0)	0.806	0.752	-0.057
	MA40_282	36	8	8.000	2 (0)	0.972	0.808	-0.190
	MA41_373	36	7	7.000	3 (0)	0.694	0.705	0.029
	MA42_001	36	5	5.000	3 (2)	0.806	0.672	-0.186
	MA42_059	36	2	2.000	0 (0)	0.389	0.346	-0.111
	MA42_126	36	2	2.000	0 (0)	0.472	0.497	0.063
	MA42_203	36	5	5.000	2 (0)	0.750	0.585	-0.268*
	MA42_231	36	5	5.000	1 (0)	0.722	0.752	0.053
	MA42_241	36	3	3.000	1 (1)	0.472	0.421	-0.108
	MA42_255	36	4	4.000	0 (0)	0.722	0.680	-0.048
	MA42_293	36	2	2.000	0 (0)	0.472	0.389	-0.202
	MA42_397	36	15	15.000	7 (0)	0.889	0.861	-0.018
	MA42_421	36	5	5.000	3 (0)	0.500	0.495	0.004
	MA42_472	36	10	10.000	2 (0)	0.750	0.834	0.114*
	Pop	36	5.250	5.250	25 (4)	0.627	0.587	-0.054
	SE	0	0.864	0.864		0.058	0.055	
Magnolia domingensis Loma Rodríguez (ROD) N_0 = 50 P = 16	MA39_023	50	4	3.720	1 (0)	0.700	0.670	-0.035
	MA39_165	50	2	2.000	0 (0)	0.360	0.403	0.117
	MA39_185	50	5	4.420	0 (0)	0.660	0.542	-0.209
	MA40_282	50	7	6.718	1 (0)	0.820	0.757	-0.074
	MA41_373	50	6	5.715	2 (1)	0.760	0.740	-0.017
	MA42_001	50	2	2.000	0 (0)	0.240	0.269	0.117
	MA42_059	50	2	1.924	0 (0)	0.040	0.039	-0.010
	MA42_126	50	2	2.000	0 (0)	0.460	0.455	-0.001
	MA42_203	50	3	3.000	0 (0)	0.660	0.615	-0.062
	MA42_231	50	6	5.720	2 (0)	0.660	0.733	0.110
	MA42_241	50	3	2.720	1 (1)	0.460	0.399	-0.142
	MA42_255	50	4	3.980	0 (0)	0.600	0.621	0.045
	MA42_293	50	2	2.000	0 (0)	0.100	0.13	0.241
	MA42_397	50	8	7.362	0 (0)	0.800	0.746	-0.062
	MA42_421	50	2	2.000	0 (0)	0.420	0.466	0.109
	MA42_472	50	9	8.567	1 (0)	0.860	0.792	-0.076
	Pop	50	4.188	3.990	8 (2)	0.538	0.524	-0.016
	SE	0	0.593	0.550		0.063	0.058	

Population	SSR marker	Summary statistics						
		N	*A*	$A_{R\text{-}36}$	A_{PA}	H_O	H_E	F_{IS}
Magnolia hamorii Bahoruco (BAH) N_0 = 118 P = 15	MA39_023	118	2	2.000	0	0.432	0.474	0.092
	MA39_165	118	1	1.000	0	0.000	0.000	M
	MA39_185	118	7	5.896	1	0.729	0.744	0.025
	MA40_282	118	17	14.959	7	0.839	0.886	0.057
	MA41_373	117	11	9.467	2	0.778	0.850	0.089
	MA42_001	118	4	3.304	1	0.568	0.569	0.007
	MA42_059	118	2	1.891	1	0.051	0.050	-0.022
	MA42_126	118	5	4.178	2	0.568	0.555	-0.018*
	MA42_203	118	9	8.062	2	0.712	0.823	0.139*
	MA42_231	118	8	7.179	3	0.788	0.781	-0.005
	MA42_241	118	2	2.000	0	0.466	0.481	0.035
	MA42_255	118	10	8.767	5	0.831	0.833	0.007
	MA42_293	118	11	9.205	6	0.788	0.830	0.055
	MA42_397	118	13	11.193	2	0.831	0.872	0.052
	MA42_421	118	2	1.518	1	0.017	0.017	-0.004
	MA42_472	118	25	19.130	13	0.873	0.916	0.051
	Pop	118	8.063	6.859	46	0.579	0.605	0.047
	SE	0	1.629	1.304		0.077	0.081	

Population	SSR marker	Summary statistics						
		N	*A*	A_{R-36}	A_{PP} (A_{PA})	H_O	H_E	F_{IS}
Magnolia pallescens Valle Nuevo (VAL) N_0 = 133 P = 16	MA39_023	132	4	3.797	1 (0)	0.371	0.359	-0.031
	MA39_165	131	2	2.000	0 (0)	0.321	0.345	0.074
	MA39_185	133	9	6.891	0 (1)	0.624	0.599	-0.038
	MA40_282	131	9	7.073	3 (1)	0.748	0.784	0.050
	MA41_373	131	5	4.095	1 (0)	0.649	0.662	0.024*
	MA42_001	131	3	2.475	0 (0)	0.389	0.489	0.208
	MA42_059	129	3	2.662	0 (1)	0.031	0.090	0.656
	MA42_126	133	4	3.501	2 (0)	0.128	0.156	0.182
	MA42_203	133	9	7.486	4 (2)	0.692	0.723	0.047
	MA42_231	133	4	3.999	1 (0)	0.594	0.527	-0.123
	MA42_241	133	2	2.000	0 (0)	0.361	0.358	-0.006
	MA42_255	133	7	6.746	0 (0)	0.759	0.808	0.064
	MA42_293	131	4	3.001	1 (0)	0.466	0.517	0.102
	MA42_397	129	15	12.538	4 (1)	0.837	0.889	0.063
	MA42_421	131	7	5.821	4 (2)	0.740	0.705	-0.047
	MA42_472	133	13	11.790	0 (0)	0.729	0.820	0.114
	Pop	131.688	6.250	5.367	21 (6)	0.527	0.552	0.048
	SE	0.35	0.968	0.811		0.060	0.060	
Magnolia pallescens Ebano Verde (EBV) N_0 = 80 P = 15	MA39_023	80	5	4.996	0 (1)	0.525	0.545	0.043
	MA39_165	80	2	2.000	0 (0)	0.325	0.289	-0.119
	MA39_185	80	7	6.402	2 (0)	0.775	0.795	0.031
	MA40_282	80	9	7.189	3 (0)	0.500	0.556	0.107
	MA41_373	80	4	3.450	2 (0)	0.488	0.572	0.154
	MA42_001	80	2	2.000	1 (0)	0.425	0.420	-0.006*
	MA42_059	80	2	2.000	1 (0)	0.288	0.312	0.086*
	MA42_126	80	6	5.348	0 (2)	0.538	0.529	-0.010*
	MA42_203	80	8	7.865	5 (2)	0.763	0.737	-0.028*
	MA42_231	80	5	4.908	0 (0)	0.625	0.635	0.021
	MA42_241	80	1	1.000	1 (0)	0.000	0.000	M
	MA42_255	80	5	5.000	2 (0)	0.788	0.776	-0.008*
	MA42_293	80	3	3.000	2 (0)	0.425	0.443	0.048
	MA42_397	80	17	14.828	2 (0)	0.838	0.900	0.076*
	MA42_421	79	9	8.676	2 (2)	0.696	0.726	0.047
	MA42_472	80	8	7.346	5 (0)	0.738	0.756	0.031*
	Pop	79.938	5.813	5.375	28 (7)	0.546	0.562	0.035*
	SE	0.063	0.993	0.861		0.056	0.058	

Locality	SSR marker	Summary statistics						
		N	A	$A_{R\text{-}6}$	A_{PP} (A_{PA})	H_O	H_E	F_{IS}
Magnolia domingensis Loma Barbacoa (BAR) N_0 = 36 P = 16	MA39_023	36	4	3.051	1 (1)	0.583	0.568	-0.014
	MA39_165	36	2	1.167	0 (0)	0.028	0.027	0.000
	MA39_185	36	5	4.340	0 (0)	0.806	0.752	-0.057
	MA40_282	36	8	5.257	2 (0)	0.972	0.808	-0.190
	MA41_373	36	7	4.495	3 (0)	0.694	0.705	0.029
	MA42_001	36	5	3.778	3 (2)	0.806	0.672	-0.186
	MA42_059	36	2	1.964	0 (0)	0.389	0.346	-0.111
	MA42_126	36	2	2.000	0 (0)	0.472	0.497	0.063
	MA42_203	36	5	3.247	2 (0)	0.750	0.585	-0.268*
	MA42_231	36	5	4.431	1 (0)	0.722	0.752	0.053
	MA42_241	36	3	2.153	1 (1)	0.472	0.421	-0.108
	MA42_255	36	4	3.657	0 (0)	0.722	0.680	-0.048
	MA42_293	36	2	1.983	0 (0)	0.472	0.389	-0.202
	MA42_397	36	15	6.937	7 (0)	0.889	0.861	-0.018
	MA42_421	36	5	2.768	3 (0)	0.500	0.495	0.004
	MA42_472	36	10	6.125	2 (0)	0.750	0.834	0.114*
	Pop	36	5.250	3.585	25 (4)	0.627	0.587	-0.054
	SE	0	0.864	0.405		0.058	0.055	
Magnolia domingensis Loma Rodríguez (ROD) N_0 = 50 P = 16	MA39_023	50	4	3.100	1 (0)	0.700	0.670	-0.035
	MA39_165	50	2	1.985	0 (0)	0.360	0.403	0.117
	MA39_185	50	5	2.560	0 (0)	0.660	0.542	-0.209
	MA40_282	50	7	4.703	1 (0)	0.820	0.757	-0.074
	MA41_373	50	6	4.178	2 (1)	0.760	0.740	-0.017
	MA42_001	50	2	1.892	0 (0)	0.240	0.269	0.117
	MA42_059	50	2	1.227	0 (0)	0.040	0.039	-0.010
	MA42_126	50	2	1.996	0 (0)	0.460	0.455	-0.001
	MA42_203	50	3	2.903	0 (0)	0.660	0.615	-0.062
	MA42_231	50	6	4.269	2 (0)	0.660	0.733	0.110
	MA42_241	50	3	2.099	1 (1)	0.460	0.399	-0.142
	MA42_255	50	4	3.214	0 (0)	0.600	0.621	0.045
	MA42_293	50	2	1.603	0 (0)	0.100	0.13	0.241
	MA42_397	50	8	4.478	0 (0)	0.800	0.746	-0.062
	MA42_421	50	2	1.997	0 (0)	0.420	0.466	0.109
	MA42_472	50	9	5.352	1 (0)	0.860	0.792	-0.076
	Pop	50	4.188	2.972	8 (2)	0.538	0.524	-0.016
	SE	0	0.593	0.316		0.063	0.058	

Locality	SSR marker	Summary statistics						
		N	A	A_{R-6}	A_{PP} (A_{PA})	H_O	H_E	F_{IS}
Magnolia hamorii Cachote (CAC) N_0 = 52 P = 13	MA39_023	52	2	1.982	0 (0)	0.462	0.393	-0.163
	MA39_165	52	1	1.000	0 (0)	0.000	0.000	M
	MA39_185	52	7	3.874	1 (0)	0.673	0.713	0.065
	MA40_282	52	16	6.775	2 (2)	0.827	0.851	0.038
	MA41_373	52	10	6.203	2 (2)	0.769	0.833	0.086
	MA42_001	52	4	2.702	1 (0)	0.596	0.565	-0.046
	MA42_059	52	1	1.000	0 (0)	0.000	0.000	M
	MA42_126	52	4	2.579	0 (0)	0.654	0.539	-0.205
	MA42_203	52	8	4.818	0 (0)	0.712	0.719	0.021
	MA42_231	52	7	4.982	0 (0)	0.788	0.791	0.013
	MA42_241	52	2	1.987	0 (0)	0.423	0.411	-0.021
	MA42_255	52	9	5.649	0 (0)	0.827	0.819	0.000
	MA42_293	52	11	5.744	3 (2)	0.750	0.819	0.094*
	MA42_397	52	10	6.389	0 (0)	0.885	0.859	-0.020
	MA42_421	52	1	1.000	0 (0)	0.000	0.000	M
	MA42_472	52	20	7.743	2 (1)	0.827	0.893	0.084
	Pop	52	7.063	4.027	11 (7)	0.575	0.575	0.011
	SE	0	1.392	0.572		0.078	0.081	
Magnolia hamorii Cortico (COR) N_0 = 52 P = 14	MA39_023	52	2	2.000	0 (0)	0.385	0.499	0.239
	MA39_165	52	1	1.000	0 (0)	0.000	0.000	M
	MA39_185	52	6	4.090	0 (0)	0.769	0.722	-0.055
	MA40_282	52	14	7.388	1 (1)	0.827	0.889	0.080
	MA41_373	52	9	5.818	1 (0)	0.808	0.823	0.028
	MA42_001	52	3	2.734	0 (0)	0.500	0.511	0.032
	MA42_059	52	2	1.392	0 (0)	0.077	0.074	-0.030
	MA42_126	52	4	2.593	0 (0)	0.481	0.473	-0.008
	MA42_203	52	8	5.424	0 (0)	0.731	0.816	0.114*
	MA42_231	52	7	4.859	1 (0)	0.750	0.743	0.000
	MA42_241	52	2	2.000	0 (0)	0.500	0.499	0.008
	MA42_255	52	9	5.552	1 (1)	0.846	0.831	-0.009
	MA42_293	52	8	5.844	0 (0)	0.808	0.820	0.025
	MA42_397	52	12	6.363	2 (1)	0.846	0.832	-0.007
	MA42_421	52	1	1.000	0 (0)	0.000	0.000	M
	MA42_472	52	18	8.011	2 (2)	0.923	0.903	-0.013
	Pop	52	6.625	4.129	8 (5)	0.578	0.590	0.029
	SE	0	1.251	0.582		0.079	0.079	

Locality	SSR marker	Summary statistics						
		N	*A*	$A_{R\text{-}6}$	A_{PP} (A_{PA})	H_O	H_E	F_{IS}
Magnolia hamorii Laguneta (LAG) $N_0 = 7$ P = 14	MA39_023	7	2	2.000	0 (0)	0.571	0.408	-0.333
	MA39_165	7	1	1.000	0 (0)	0.000	0.000	M
	MA39_185	7	3	2.989	0 (0)	0.857	0.602	-0.358
	MA40_282	7	10	8.978	0 (0)	0.857	0.878	0.100
	MA41_373	7	3	3.000	0 (0)	0.571	0.663	0.213
	MA42_001	7	2	2.000	0 (0)	0.714	0.459	-0.500
	MA42_059	7	2	1.989	0 (0)	0.286	0.245	-0.091
	MA42_126	7	4	3.714	0 (0)	0.571	0.622	0.158
	MA42_203	7	5	4.703	1 (1)	0.714	0.704	0.063
	MA42_231	7	5	4.835	0 (0)	1.000	0.724	-0.313
	MA42_241	7	2	2.000	0 (0)	0.714	0.500	-0.364
	MA42_255	7	4	3.857	0 (0)	0.714	0.694	0.048
	MA42_293	7	4	3.857	0 (0)	0.857	0.694	-0.161
	MA42_397	7	4	3.846	0 (0)	0.571	0.643	0.186
	MA42_421	7	1	1.000	0 (0)	0.000	0.000	M
	MA42_472	7	8	7.264	0 (0)	0.857	0.806	0.014
	Pop	7	3.750	3.565	1 (1)	0.616	0.540	-0.064
	SE	0	0.609	0.539		0.073	0.065	
Magnolia hamorii Tonobán (TON) $N_0 = 7$ P = 14	MA39_023	7	2	2.000	0 (0)	0.429	0.459	0.143
	MA39_165	7	1	1.000	0 (0)	0.000	0.000	M
	MA39_185	7	4	4.000	0 (0)	0.714	0.745	0.118
	MA40_282	7	10	8.978	0 (0)	1.000	0.878	-0.063
	MA41_373	6	6	6.000	0 (0)	0.833	0.750	-0.020
	MA42_001	7	3	2.857	0 (0)	0.714	0.541	-0.250
	MA42_059	7	1	1.000	0 (0)	0.000	0.000	M
	MA42_126	7	4	3.714	1 (0)	0.571	0.531	0
	MA42_203	7	5	4.835	0 (0)	0.571	0.745	0.304
	MA42_231	7	5	4.703	0 (0)	0.857	0.704	-0.143
	MA42_241	7	2	2.000	0 (0)	0.286	0.490	0.478
	MA42_255	7	6	5.560	0 (0)	0.857	0.735	-0.091
	MA42_293	7	5	4.714	0 (0)	0.857	0.735	-0.091
	MA42_397	7	6	5.692	0 (0)	0.571	0.776	0.333
	MA42_421	7	2	1.989	1 (1)	0.286	0.245	-0.091
	MA42_472	7	8	7.407	2 (2)	0.857	0.847	0.065
	Pop	6.938	4.375	4.153	4 (3)	0.588	0.574	0.053
	SE	0.063	0.632	0.572		0.078	0.069	

Locality	SSR marker	Summary statistics						
		N	*A*	A_{R-6}	A_{PP} (A_{PA})	H_O	H_E	F_{IS}
Magnolia pallescens Casabito (CAS) N_0 = 40 P = 15	MA39_023	40	5	3.334	0 (0)	0.525	0.480	-0.081
	MA39_165	40	2	1.933	0 (0)	0.325	0.305	-0.054
	MA39_185	40	6	4.487	0 (0)	0.800	0.751	-0.052
	MA40_282	40	7	3.538	1 (0)	0.525	0.534	0.030
	MA41_373	40	3	2.631	0 (0)	0.475	0.535	0.125
	MA42_001	40	2	1.997	0 (0)	0.425	0.462	0.093
	MA42_059	40	2	1.977	0 (0)	0.350	0.375	0.079
	MA42_126	40	5	3.095	0 (0)	0.550	0.529	-0.028
	MA42_203	40	7	4.522	1 (1)	0.775	0.744	-0.029
	MA42_231	40	5	3.616	1 (0)	0.650	0.552	-0.165
	MA42_241	40	1	1.000	0 (0)	0.000	0.000	M
	MA42_255	40	5	4.406	0 (0)	0.725	0.746	0.041
	MA42_293	40	3	2.730	0 (0)	0.475	0.524	0.106
	MA42_397	40	13	6.653	1 (0)	0.800	0.848	0.069
	MA42_421	40	8	3.282	0 (0)	0.475	0.476	0.014
	MA42_472	40	8	5.022	0 (0)	0.850	0.788	-0.065
	Pop	40	5.125	3.389	4 (1)	0.545	0.541	0.004
	SE	0	0.769	0.352		0.055	0.053	
Magnolia pallescens Entrance (ENT) N_0 = 61 P = 16	MA39_023	61	4	2.143	0 (0)	0.230	0.224	-0.016
	MA39_165	61	2	1.952	0 (0)	0.295	0.335	0.128
	MA39_185	61	6	2.912	0 (0)	0.574	0.489	-0.166
	MA40_282	61	5	4.236	0 (0)	0.754	0.723	-0.035
	MA41_373	61	3	2.965	0 (0)	0.672	0.648	-0.030
	MA42_001	61	2	1.995	0 (0)	0.426	0.451	0.064
	MA42_059	59	3	1.865	1 (1)	0.051	0.173	0.710*
	MA42_126	61	4	1.784	0 (0)	0.131	0.140	0.071
	MA42_203	61	8	4.477	1 (0)	0.672	0.723	0.079*
	MA42_231	61	4	3.466	0 (0)	0.705	0.603	-0.162
	MA42_241	61	2	1.945	0 (0)	0.377	0.326	-0.149
	MA42_255	61	6	4.438	0 (0)	0.639	0.756	0.162*
	MA42_293	61	4	2.282	1 (0)	0.426	0.487	0.134
	MA42_397	58	13	6.791	0 (0)	0.776	0.866	0.113*
	MA42_421	61	4	3.730	0 (0)	0.738	0.682	-0.074
	MA42_472	61	13	6.598	1 (0)	0.852	0.836	-0.011
	Pop	60.688	5.188	3.349	4 (1)	0.520	0.529	0.025*
	SE	0.218	0.862	0.404		0.062	0.059	

Locality	SSR marker	Summary statistics						
		N	*A*	A_{R-6}	A_{PP} (A_{PA})	H_O	H_E	F_{IS}
Magnolia pallescens Camino Guayabal (GUA) N_0 = 15 P = 15	MA39_023	14	4	2.540	0 (0)	0.286	0.258	-0.072
	MA39_165	14	2	1.911	0 (0)	0.286	0.245	-0.130
	MA39_185	15	6	3.985	1 (1)	0.667	0.527	-0.233
	MA40_282	14	3	2.827	0 (0)	0.500	0.559	0.142
	MA41_373	14	4	3.51	1 (0)	0.714	0.630	-0.097
	MA42_001	14	2	1.996	0 (0)	0.286	0.408	0.333
	MA42_059	14	1	1.000	0 (0)	0.000	0.000	M
	MA42_126	15	3	2.384	0 (0)	0.400	0.407	0.051
	MA42_203	15	6	3.939	0 (0)	0.667	0.558	-0.162
	MA42_231	15	4	3.185	0 (0)	0.533	0.436	-0.191
	MA42_241	15	2	1.984	0 (0)	0.467	0.358	-0.273
	MA42_255	15	5	4.571	0 (0)	0.733	0.753	0.061
	MA42_293	14	2	2.000	0 (0)	0.357	0.497	0.316
	MA42_397	15	10	5.831	0 (0)	0.867	0.764	-0.100
	MA42_421	14	4	3.806	0 (0)	0.857	0.712	-0.169
	MA42_472	15	7	4.695	0 (0)	0.667	0.609	-0.061
	Pop	14.500	4.063	3.135	2 (1)	0.518	0.482	-0.038
	SE	0.129	0.581	0.321		0.060	0.051	
Magnolia pallescens Loma de la Sal (SAL) N_0 = 40 P = 15	MA39_023	40	5	3.902	0 (0)	0.525	0.597	0.133
	MA39_165	40	2	1.900	0 (0)	0.325	0.272	-0.182
	MA39_185	40	7	5.018	0 (0)	0.750	0.798	0.072
	MA40_282	40	8	3.551	1 (0)	0.475	0.572	0.182
	MA41_373	40	4	3.039	0 (0)	0.500	0.582	0.154*
	MA42_001	40	2	1.971	0 (0)	0.425	0.362	-0.161
	MA42_059	40	2	1.854	0 (0)	0.225	0.237	0.064
	MA42_126	40	6	3.583	1 (1)	0.525	0.515	-0.007
	MA42_203	40	6	4.375	1 (0)	0.750	0.669	-0.108
	MA42_231	40	4	3.388	0 (0)	0.600	0.650	0.089
	MA42_241	40	1	1.000	0 (0)	0.000	0.000	M
	MA42_255	40	5	4.634	0 (0)	0.850	0.786	-0.069
	MA42_293	40	3	2.269	0 (0)	0.375	0.316	-0.175
	MA42_397	40	16	7.795	0 (0)	0.875	0.901	0.041
	MA42_421	39	9	6.299	1 (1)	0.923	0.851	-0.071
	MA42_472	40	7	4.400	0 (0)	0.625	0.667	0.076
	Pop	39.938	5.438	3.686	4 (2)	0.547	0.548	0.016
	SE	0.063	0.917	0.441		0.062	0.063	

Locality	SSR marker	Summary statistics						
		N	*A*	$A_{R\text{-}6}$	A_{PP} (A_{PA})	H_O	H_E	F_{IS}
Magnolia pallescens Montellano (MON) N_0 = 41 P = 15	MA39_023	41	4	2.925	0 (0)	0.561	0.505	-0.099
	MA39_165	41	2	1.960	0 (0)	0.341	0.343	0.016
	MA39_185	41	7	3.924	0 (0)	0.707	0.667	-0.047
	MA40_282	41	9	5.269	1 (1)	0.805	0.792	-0.003
	MA41_373	41	4	3.236	0 (0)	0.610	0.662	0.090
	MA42_001	41	2	2.000	0 (0)	0.390	0.493	0.220
	MA42_059	41	1	1.000	0 (0)	0.000	0.000	M
	MA42_126	41	3	1.293	0 (0)	0.049	0.048	-0.006
	MA42_203	41	7	4.191	1 (1)	0.707	0.690	-0.012
	MA42_231	41	4	2.607	0 (0)	0.390	0.336	-0.150
	MA42_241	41	2	1.928	0 (0)	0.268	0.299	0.115
	MA42_255	41	6	4.720	0 (0)	0.854	0.790	-0.068
	MA42_293	41	3	2.146	0 (0)	0.610	0.511	-0.181
	MA42_397	41	13	7.131	0 (0)	0.902	0.884	-0.009
	MA42_421	41	7	4.601	1 (1)	0.634	0.726	0.139
	MA42_472	41	9	5.993	0 (0)	0.610	0.843	0.288*
	Pop	41	5.188	3.433	3 (3)	0.527	0.537	0.030*
	SE	0	0.823	0.443		0.067	0.068	
Magnolia pallescens Rancho Guaraguao (RAN) N_0 = 16 P = 16	MA39_023	16	4	2.949	0 (0)	0.500	0.443	-0.096
	MA39_165	15	2	1.999	0 (0)	0.400	0.444	0.134
	MA39_185	16	5	3.366	0 (0)	0.563	0.604	0.100
	MA40_282	15	4	3.776	0 (0)	0.800	0.704	-0.102
	MA41_373	15	4	3.288	0 (0)	0.600	0.629	0.080
	MA42_001	15	3	2.645	1 (0)	0.333	0.504	0.369*
	MA42_059	15	2	1.400	0 (0)	0.067	0.064	0.000
	MA42_126	16	3	1.992	0 (0)	0.063	0.174	0.659*
	MA42_203	16	7	5.468	0 (0)	0.750	0.789	0.082
	MA42_231	16	4	3.551	0 (0)	0.750	0.631	-0.158
	MA42_241	16	2	2.000	0 (0)	0.438	0.498	0.153
	MA42_255	16	6	5.000	0 (0)	1.000	0.785	-0.244
	MA42_293	15	3	2.369	0 (0)	0.333	0.371	0.136
	MA42_397	15	10	7.155	0 (0)	0.867	0.864	0.032
	MA42_421	15	5	3.598	0 (0)	0.933	0.627	-0.463*
	MA42_472	16	8	5.389	0 (0)	0.625	0.727	0.171
	Pop	15.500	4.500	3.496	1 (0)	0.564	0.554	0.015
	SE	0.129	0.570	0.391		0.071	0.055	