Podocarpus in the palaeogeographically complex island of Hispaniola: A stepping-stone colonization and conservation recommendations

María Esther Nieto-Blázquez1,2 | María Paula Quiroga3 | Andrea C. Premoli3 | Julissa Roncal1

Abstract
Aim: Hispaniola is the second largest island in the Caribbean and a hot spot of biodiversity. The island was formed by the fusion of a northern and southern palaeo-islands during the mid-Miocene (15 Ma). The historical split of Hispaniola together with repeated marine incursions during the Pleistocene is known to have influenced lineage divergence and genetic structure in a few birds and mammals, but the effect on vascular plants is less understood. The conifer genus Podocarpus has two species, P. hispaniolensis and P. buchii, that are endemic to the mountains of Hispaniola and are IUCN endangered. The former occurs in the mountains of the north, and the latter in the south, with a region of sympathy in the Cordillera Central. Here, we evaluate the historical split of the two palaeo-islands and repeated marine incursions as dispersal barriers to the geographical distribution of genetic diversity, genetic structure, divergence patterns and the historical demography of the two species.

Location: Hispaniola island, Caribbean.

Methods: Using genotyping-by-sequencing in 47 Podocarpus samples, we identified two sets of single nucleotide polymorphisms for our analyses (74,260 and 22,657 SNPs). We conducted a phylogenetic and an approximate Bayesian computation analysis to test evolutionary hypotheses of sympatric and allopatric speciation and stepping-stone colonization.

Results: Podocarpus showed a population genetic structure that corresponds to the geographic distribution of the species. Podocarpus on Hispaniola fit a stepping-stone colonization model with bottlenecks at each mountain colonization event and speciation in Cordillera Central.

Main conclusions: The historical events in question did not seem to have influenced the genetic structure, diversity or demography of Podocarpus; instead, the current geographic barriers imposed by lowland xeric valleys did. The clear divergence between species together with the elevated within-population genetic diversity and significant genetic structure calls for a multi-population in situ conservation of each species.
**1 | INTRODUCTION**

The Caribbean archipelago is the largest, most species-rich insular system in the Neotropics (Maunder et al., 2008), and one of the world’s 34 biodiversity hot spots sensu Mittermeier et al. (2004). In situ speciation is one of the leading hypotheses proposed to explain its high biodiversity (Roncal et al., 2020) which is supported by the high number of endemics (up to 8000 seed plant species; Acevedo-Rodríguez & Strong, 2008). Endemic clades thus provide opportunities for understanding the evolutionary history of insular biotas. Previous studies have aimed at reconstructing the colonization history and adaptive radiations of Caribbean lineages with a bias for vertebrates (e.g. Losos et al., 2006; Nieto-Blázquez et al., 2021; Reynolds et al., 2016; Turvey et al., 2016). The heterogeneous environment of large and geologically complex islands has been associated with historical and modern barriers to gene flow that can drive speciation (Turvey et al., 2016). The evolutionary history of intra-island endemic plant lineages, that may or may not surpass these barriers, will thus increase our understanding of how effective these barriers are for intra-island diversification. Moreover, the distribution of genetic diversity in large islands will provide insight on gene flow in space and will inform local conservation strategies for insular taxa (Brown et al., 2014).

The island of Hispaniola is the result of the juncture of two palaeo-islands that collided tectonically during the middle Miocene, ca. 15 Ma (Graham, 2003; Mann et al., 1991). These northern and southern palaeo-islands remained separated by the Neiba Valley-Cul de Sac Plain (Figure 1), an arid and deep rift that was inundated by marine incursions repeatedly over the late Pleistocene (Maurrasse et al., 1980; McLaughlin et al., 1991). The southern palaeo-island is further divided into two physiographic provinces (Massif de la Hotte to the west and Massif de la Selle-Sierra de Bahoruco to the east) separated by the Jacmel-Fauche’ depression (known as Bond’s line), which was also inundated during the Plio-Pleistocene through a sea channel (Maurrasse et al., 1980). These geological events left three distinct biogeographic regions disconnected during the incursion periods generating allopatric populations and genetic divergence in a few endemic birds and mammals (Brace et al., 2012; Sly et al., 2010, 2011; Townsend et al., 2007; Turvey et al., 2016), but the genetic or evolutionary effects on other lineages are less understood including vascular plants and thus preventing the generalization of this model.

Hispaniola’s current topography consists of mountain chains (cordilleras) that run in parallel from NW to SE (Latta et al., 2006) separated by lowland xeric valleys (Heubeck & Mann, 1991; Townsend et al., 2007). The largest mountain system is the Cordillera Central, which was uplifted during the middle to late Eocene and is dominated
by igneous and volcanic materials. The Sierra Neiba, a karstic range in the south of Hispaniola (or southern range of the northern palaeo-island), and the Sierra Bahoruco, a limestone mountain in the south-central part of Hispaniola, are of similar age to the Cordillera Central. The Cordillera Septentrional, located in the north of Hispaniola (or northern range of the northern palaeo-island), is of later, Oligocene–Miocene origin and presents mostly sedimentary rocks (Figure 1; Cano-Carmona et al., 2010; Cano-Ortiz et al., 2016). These mountain chains have been invoked as the explanation for the high levels of genetic differentiation among populations of a dry lowland lizard species (Gifford et al., 2004). Other current ecological factors that have caused lineage divergence of Hispaniolan populations are geographic macrohabitats (in lizards, Glor et al., 2003) and unoccupied feeding niches (e.g. in butterflies, Matos-Maravi et al., 2014).

The conifer genus Podocarpus is represented by two endemic sister species on Hispaniola, *P.* buchii Urb. and *P.* hispaniolensis de Laub., that diverged 15 mya (Nieto-Blázquez et al., 2021), and coinciding with the junction of the two palaeo-islands. *Podocarpus* colonized the Greater Antilles from the Andes 45 to 31 mya, and a vicariant origin for the ancestor of Hispaniolan *Podocarpus* attributed to the tectonic separation of eastern Cuba and western Hispaniola during the middle Miocene has been proposed (Nieto-Blázquez et al., 2021). The sister species differ in morphology, elevational range and distribution. *Podocarpus buchii* can be distinguished from *P. hispaniolensis* by the absence of a prominent deep grove on the adaxial leaf surface, the smaller terminal buds usually less than 3.5 mm (not >5 mm), the short female peduncles less than 4 mm long (not >6 mm) and the smaller male cones less than 10 mm (not >12 mm, Mill, 2015). *Podocarpus buchii* tends to be found at higher elevations (1100–2500 m) than *P. hispaniolensis* (800–1200 m). *Podocarpus hispaniolensis* is found in the northernmost mountain chain (Cordillera Septentrional) of the Dominican Republic and historically in the Massif du Nord in Haiti (Mill, 2015). *Podocarpus buchii* is found in the south, particularly in the Sierra de Bahoruco and Neiba of the Dominican Republic, and also in the Massif de la Hotte and Chaîne de la Selle in Haiti (Mill, 2015, Figure 1). Both species occur in the Cordillera Central, and although the two species do not typically occur in sympathy, collections that include both species have been made in relatively small areas. Seeds are probably dispersed by frugivorous birds and small mammals attracted by the colour and/or swollen cone bract (Enright & Jaffré, 2011; Mill, 2003). Several bird families feed on the fleshy female cones of the congenic *P. parlatorei* Pilg. transporting its seeds (Blendinger, 2017).

Both Hispaniolan species are IUCN endangered because of habitat loss due to logging, forest clearance for agricultural and commercial plantations, fire and mining, and selective logging of its highly valuable timber (Gardner, 2013). The most severe habitat degradation and loss has occurred in Haiti. The trend of population decline for *P. buchii* has been estimated at 50% over three generations, resulting in currently severely fragmented populations (Gardner, 2013). *Podocarpus hispaniolensis* has an estimated area of occupancy of less than 5000 km² with about five locations severely fragmented. Both species are recorded in a few protected areas where threats are still present. There is no action recovery plan, and no local species management including ex situ conservation for either species (Gardner, 2013).

Vulnerable and endangered species usually consist of small and relatively isolated populations that maybe impacted by the detrimental effects of reduced genetic diversity due to drift, limited gene flow and inbreeding that can threaten their long-term persistence (Frankham et al., 2002). Although the impact of genetic factors on species’ well-being has been controversial (Spelman et al., 2004; Teixeira & Huber, 2021), populations occurring in small numbers should be the focus of conservation actions because they may be prone to suffer from environmental and demographic stochasticity (Lande, 1993), the latter usually driven by anthropogenic factors. Nonetheless, with the advent of the genomic era, we have a unique opportunity to analyse in more detail populations and species under threat in order to contribute to their better conservation and management (Supple & Shapiro, 2018).

Our study aims to advance our knowledge of whether multiple Hispaniolan lineages exhibit similar evolutionary history and population structure in response to historical allopatric barriers (e.g. palaeo-islands, marine incursions), or whether the significance of these historical barriers in regulating gene flow is now masked by current landscape factors (e.g. topography, lowland xeric valleys). Our research (1) estimates genetic diversity of populations, (2) identifies genetic structure within and between species, (3) infers the phylogenetic relationship among populations, (4) elucidates the demographic history of the genus on the island and (5) recommends strategies of long-term genomic conservation of Hispaniolan *Podocarpus*. We tested four alternative hypotheses for analysing the historical demography: (1) an ancestral population from Cordillera Central gave rise to both *Podocarpus* species from where one species migrated south and the other north, (2) allopatric speciation in the two palaeo-islands followed by secondary contact in Cordillera Central, (3) a stepping-stone colonization from south to north and (4) a stepping-stone colonization from north to south.

2 MATERIAL AND METHODS

2.1 Sample collection and DNA extraction

We collected silica-dried leaf samples from 20 *P. buchii* and 16 *P. hispaniolensis*, across 11 collection sites in the Dominican Republic across all four mountain chains (Table 1, Figure 1). The low sample size reflects their rarity in the field, as recognized by their status as IUCN Endangered and the alarming estimate that forest cover in the Dominican Republic is down to 14% (Russell, 1991). Very few individuals were thus available for sampling. Duplicate herbarium specimens from each individual were deposited at the National Herbarium of the Jardín Botánico Nacional Dr. Rafael Ma. Moscoco (UBSD) and the Royal Botanic Garden of Edinburgh (E, Table S1). We selected adult individuals with healthy-looking leaves for both silica-dried leaf samples and herbarium specimens. We selected as out-group
TABLE 1 Sample collection information for Podocarpus on Hispaniola

| Species               | Collection Site       | Locality                     | Latitude   | Longitude   | Elevation (in metres) | # indiv |
|-----------------------|-----------------------|------------------------------|------------|-------------|-----------------------|---------|
| *Podocarpus buchii*   | 1 (SB-b)              | Barahona Province, Sierra Bahoruco Oriental | 18°0.915′N | 71°17.017′W | 1474                  | 5       |
|                       | 2 (SB-b)              | Pedernales Province, Sierra Bahoruco (DR) Occidental (DR) | 18°12.495′N | 71°33.353′W | 1525                  | 1       |
|                       | 3 (SN-b)*             | Loma del Quince, Independencia Province, Sierra de Neiba (DR) | 18°40.515′N | 71°41.182′W | 1855                  | 5       |
|                       | 4 (SN-b)              | Loma del Quince, Independencia Province, Sierra de Neiba (DR) | 18°41.478′N | 71°47.258′W | 1966                  | 5       |
|                       | 5 (CC-b)              | Constanza area, La Vega Province, Cordillera Central (DR) | 18°58.040′N | 70°47.770′W | 1728                  | 2       |
|                       | 6 (CC-b)              | Constanza area, La Vega Province, Cordillera Central (DR) | 18°43.565′N | 70°52.255′W | 1386                  | 2       |
| *Podocarpus hispaniolensis* | 5 (CC-h)*            | San Jose de Ocoa Province, Rancho Arriba, Cordillera Central (DR) | 18°40.193′N | 70°21.792′W | 734                   | 4       |
|                       | 6 (CC-h)*             | San Jose de Ocoa Province, Loma del Filme, Cordillera Central (DR) | 18°43.687′N | 70°22.530′W | 1103                  | 2       |
|                       | 9 (CC-h)*             | Ebano Verde Scientific Reserve, Casabito, La Vega Province, Cordillera Central (DR) | 19°2.28′N | 70°31.08′W | 1500                  | 1       |
|                       | 10 (CS-h)             | Yaroa, Puerto Plata Province, Cordillera Septentrional (DR) | 19°35.283′N | 70°36.047′W | 783                   | 5       |
|                       | 11 (CS-h)             | Tenares, La Jibara, Salcedo Province, Cordillera Septentrional (DR) | 19°29.605′N | 70°20.005′W | 733                   | 4       |

Note: Collection site numbers as in Figure 1, with acronyms in parenthesis indicating the designated population for historical demographic analysis in DIYABC. # indiv = number of individuals collected per site. Individuals from collection sites marked with an * occur in protected areas (Figure S5). Abbreviation: DR, Dominican Republic.

single specimens from 11 of the 13 Podocarpus species in the tropical American clade to which Hispaniola’s Podocarpus belongs (P. angustifolius, P. aristulatus, P. coriaceus, P. ekmanii, P. guatemalensis, P. matudae, P. oleifolius, P. sellowii, P. trinitensis, P. urbanii and P. coriaceus l.); the time of incubation for cell lysis up to 60 min and P3 buffer (for precipitation of polysaccharides, detergent, and proteins) from 195 to 225 μl.

Library preparation and sequencing followed the protocol of Abed et al. (2019). We obtained single-end sequencing reads of variable length (up to 200 bp) from two chips of an Ion Proton system. We inspected data quality with FastQC 0.11.5 (Babraham Bioinformatics), which reports Phred quality scores and % GC content.

We conducted SNP discovery with Stacks 1.47 (Catchen et al., 2013). We demultiplexed and filtered sequencing reads using the process_radtags function. We trimmed reads to 64 bp length and eliminated uncalled reads and reads with low-quality scores (phred score of 10). Since we did not have a reference genome, we conducted a “de novo” assembly of loci with the module “ustacks.” In this step, reads were aligned into matching blocks or stacks per sample. Then, the stacks were compared, a set of putative loci produced and SNPs detected at each locus. We followed the recommendations of Paris et al. (2017) on the selection of parameter values through the pipeline, keeping those that rendered the highest number of loci. We used -m (minimum depth coverage to create a stack) = 5; -M (maximum distance allowed to create a stack) = 3; -p (parallel execution of several threads) = 15; and the -gapped option which allowed gapped alignments between stacks. Subsequently, the module “cstacks” built a catalog from the loci produced in “ustacks” by merging stacks with at most 3 (n = 3) mismatches between loci and allowed parallel execution of several threads (-p = 15). The module “sstacks” matches the loci produced by “ustacks” with the catalog produced by “cstacks,” and

2.2 | Genotyping-by-sequencing, SNP discovery and genetic diversity

The Institut de Biologie Intégrative et de Systèmes (IBIS) of the University of Laval in Canada conducted the genotyping-by-sequencing (GBS). Genomic libraries were prepared for the 47 DNA samples with two restriction enzymes, SbfI (high fidelity) andMspI (New England BioLabs Inc.). Unique barcodes 10–12 bp long were added to each sample to facilitate subsequent demultiplexing. The Institut de Biologie Intégrative et de Systèmes (IBIS) of the University of Laval in Canada conducted the genotyping-by-sequencing (GBS). Genomic libraries were prepared for the 47 DNA samples with two restriction enzymes, SbfI (high fidelity) and MspI (New England BioLabs Inc.). Unique barcodes 10–12 bp long were added to each sample to facilitate subsequent demultiplexing.
again parallel execution of several threads (\( p = 15 \)) was allowed. We used the “populations” module of Stacks to obtain population-level summary statistics, genetic diversity indexes (i.e. observed and expected heterozygosity, nucleotide polymorphism diversity), \( F_{ST} \) values and a SNP dataset. We chose the following parameters: \( -p \) (minimum number of populations containing each locus) = 1; \( -r \) (minimum percentage of individuals in a population to process a locus for that population) = 0.4; and \( -\text{min}_\text{maf} \) (minimum allele frequency) = 0.1. We ran the “populations” module of Stacks twice. In the first run, we assigned samples to the 11 collection sites plus the out-group, and in the second, we assigned samples to five populations (see Section 3). The first SNP dataset was used to run a phylogenetic reconstruction, while the second to run the genetic structure, and demographic history analyses described below.

We performed a Mantel test to detect isolation by distance (IBD). We used the R package “ade4” (Chessel et al., 2004) to measure the correlation between a matrix of pairwise \( F_{ST} \) values among the 11 collection sites and their geographic distances. We calculated pairwise geographic distances in metres with the Geographic Distance Matrix Generator (Ersts, 2011). The mantel test uses the Pearson coefficient and a default alpha value of 0.05.

2.3 | Genetic diversity and structure of Podocarpus on Hispaniola

To identify loci under selection, we performed an outlier analysis using the Bayesian-likehood method implemented in BayeScan 2.1 (Foll & Gaggiotti, 2008) using the following parameters: number of outputted iterations was \( 5 \times 10^3 \), thinning interval size of 10, 20 pilot runs of length \( 5 \times 10^3 \) each and burn-in length of 50 \( \times 10^3 \). We applied a false discovery rate of 0.05.

Genetic diversity statistics were calculated for the 11 Podocarpus collection sites on Hispaniola island by means of \( H_o = \) observed heterozygosity, \( H_e = \) expected heterozygosity, \( x = \) nucleotide polymorphism and number of private, that is unique, alleles. We investigated the genetic structure among individuals with a model-based Bayesian clustering method as implemented in STRUCTURE v2.3.4 (Pritchard et al., 2000). We ran STRUCTURE for the two Podocarpus species separately and combined. We determined the number of genetic clusters (K) with default parameter settings, the admixture model and correlated allele frequencies. We ran analyses with and without \( a \text{ priori} \) population information (PopData) for comparison. A total of five iterations of \( K = 1-10 \) were run for 500,000 Markov chain Monte Carlo (MCMC) generations, following 100,000 burn-in generations. We ran STRUCTURE with three datasets in the combined Podocarpus analysis: all loci, neutral loci and loci under selection. We used Structure Harvester v0.6.94 (Earl & vonHoldt, 2012; available at http://taylor0.biology.ucla.edu/structureHarvester/) to visualize STRUCTURE results and to explore different likelihood values across all Ks of all iterations in the analysis. Structure Harvester provides the lnP(D(K)) and uses the Evanno method for the estimation of \( \Delta K \) (Evanno et al., 2005). Finally, we used distract OSX10 (Rosenberg, 2004) to graphically display the STRUCTURE results.

2.4 | Phylogenetic relationships

To determine the phylogenetic relationships among collection sites, we ran maximum-likelihood (ML) and Bayesian inference (BI) analyses using the CIPRESS Science Getaway (Miller et al., 2010), including all individuals per collection site. We ran the ML analysis on RaxML-HPC2 v8.2.10 (Stamatakis, 2006) where we used the nucleotide substitution model GTR+I and a rapid bootstrap (BS) algorithm with 500 replicates. We ran the BI analysis with the GTR+G substitution model on MrBayes v3.2.6 (Ronquist et al., 2012) for 50 million MCMC generations sampling every 10,000 generations, and a burn-in of 25% generations discarded. We selected 11 tropical American Podocarpus species as the out-group. To further elucidate the pattern of Podocarpus population divergence, we also ran ML and BI phylogenetic analyses for the five populations used in the historical demographic analysis using the same parameters as described above.

2.5 | Demographic history of Podocarpus on Hispaniola

To infer the demographic history of Podocarpus on Hispaniola, we conducted an approximate Bayesian computation analysis in DIYABC v 2.1.0 (Cornuet et al., 2014). We defined five populations based on the genetic structure clusters (see Section 3) and their geographic distribution: SB-b = P.buchii from Sierra Bahoruco, SN-b = P.buchii from Sierra Neiba, CC-b = P.buchii from Cordillera Central, CC-h = P.hispaniolensis from Cordillera Central and CS-h = P.hispaniolensis from Cordillera Septentrional (Figure 1). Population sample sizes ranged from 4 to 10 (Table 2). Population sizes are designated as N (e.g. NSB-b corresponds to P.buchii from Sierra Bahoruco). Population sizes during bottlenecks from founder events are designated as Nf, and population size of an unsampled ancestor is designated as NA.

We designed eight competing demographic models to represent the four different evolutionary hypotheses (Figure 2). Model 1 represents an ancestral population from Cordillera Central that gave rise to both Podocarpus species from where subsequently one species migrated south and the other north. Model 2 has identical divergence patterns to Model 1 but includes bottlenecks for each migration event. Model 3 proposes that the two species originated allopatrically in the two palaeo-islands, reaching secondary contact in the Cordillera Central. Model 4 has identical topology to Model 3 but includes bottlenecks. Model 5 proposes a stepping-stone migration pattern from the south to the north. Model 6 has identical topology to Model 5 but with bottlenecks. Model 7 proposes a stepping-stone migration pattern from the north to the south. Model 8 has identical topology to Model 7 but with bottlenecks (Figure 2). Bottlenecks in Models 2, 4, 6 and 8 were included after each colonization step to model founder effects. Additional information for the DIYABC analysis can be found in the Supplementary Material.
3 | RESULTS

3.1 | Genotyping-by-sequencing and SNP discovery

The Ion Proton sequencing generated 32.36 GB of raw data containing ca. 102.6 million reads. After demultiplexing, quality filtering and discarding reads with ambiguous barcodes, a total of ca. 93.7 million reads (91.3% of initial number of reads) were kept for further analyses. The number of reads per sample ranged from ca. 258 M to ca. 10.3 M, with an average number of ca. 2.7 M reads per sample. Read quality inspection in FastQC reported a Phred score average of 24, and GC content of 57%. The catalog built from Stacks produced 788,666 loci and 276,718 SNPs. After filtering in the Stacks populations module and assigning samples to the 11 collection sites, 116,547 loci (of which 32,456 were polymorphic) and 74,260 SNPs remained for the phylogenetic reconstruction (outgroup included). If samples were instead assigned to the five populations and excluded the out-group, 11,615 loci (of which 8335 were polymorphic) and 22,657 SNPs were recovered. BayeScan identified 48 loci under selection. SNP data matrices are available in Dryad.

3.2 | Patterns of genetic diversity and structure of \textit{Podocarpus} on Hispaniola

Population genetic statistics of observed ($H_o$) and expected heterozygosity ($H_e$) and nucleotide polymorphism diversity ($\pi$) showed that \textit{P.buchii} had higher genetic diversity than \textit{P.hispaniolensis} (Table 2). This result was consistent with the higher number of genetic clusters found in \textit{P.buchii} compared with \textit{P.hispaniolensis}. When individuals were grouped into five populations, SN-b had the highest $H_e$ followed by SB-b, while the northernmost population (CS-h) had the lowest $H_e$. CC-b and SB-b showed the highest $H_o$ and $\pi$, and CS-h showed the lowest $\pi$ (Table S2). When individuals were grouped by collection site, $H_e$ and $\pi$ were highest for the two Sierra Neiba sites (3 and 4) with \textit{P.buchii}, while the lowest values were observed on collection Site 6 for \textit{P.hispaniolensis} in Cordillera Central. Site 2 (\textit{P.buchii} in SB) had highest $H_o$ while Site 5 (\textit{P.hispaniolensis} in CC) had lowest $H_o$ (Table 1). Pairwise $F_{ST}$ values among the 11 collection sites ranged from <0.001 to 0.160 (Table S3). The Mantel test showed a significant correlation between the $F_{ST}$ matrix for the 11 collection sites and their geographic location ($R = 0.442, p$-value = 0.003), which suggests the presence of isolation by geographic distance.

The STRUCTURE analyses with and without sample assignment into populations gave very similar results, and thus, we present results with population assignments. The full dataset with both species combined contained 22,657 SNPs, of which 22,609 were neutral, and 48 were under selection. The genetic structure analysis for \textit{P.buchii} used 15,524 SNPs and for \textit{P.hispaniolensis} used 5687 SNPs. For \textit{P.buchii}, the highest supported $K$ was 4 followed by $K=2$. Both $K=2$ and $K=4$ showed individuals from Sierra Bahoruco as a distinct genetic cluster. Relative to $K=2$, $K=4$ distinguished individuals in Cordillera Central from those in Sierra de Neiba (Figure S1). Two individuals from Sierra Bahoruco and the four from Cordillera Central showed admixture for $K=2$, while individuals from Sierra Neiba did not show any admixture. For $K=4$, four individuals from Sierra Bahoruco, two from Sierra Neiba and one from Cordillera Central showed admixture (Figure S1). For \textit{P.hispaniolensis}, the highest supported number of clusters was $K=2$ followed by $K=5$. $K=2$ differentiated individuals in Cordillera Central from those in Cordillera Septentrional.

### TABLE 2 Genetic diversity statistics for the 11 \textit{Podocarpus} collection sites on Hispaniola Island

| Species                  | Collection site | $H_o$   | $H_e$   | $\pi$   | # of private alleles |
|--------------------------|-----------------|---------|---------|---------|----------------------|
| \textit{Podocarpus buchii} | 1               | 0.1904  | 0.3030  | 0.3894  | 4908                 |
|                          | 2               | 0.2532  | 0.1266  | 0.2532  | 629                  |
|                          | 3               | 0.1580  | 0.3515  | 0.4559  | 2683                 |
|                          | 4               | 0.1410  | 0.3051  | 0.3943  | 3336                 |
|                          | 7               | 0.2404  | 0.1394  | 0.2611  | 781                  |
|                          | 8               | 0.1665  | 0.1057  | 0.1888  | 824                  |
| Average                  |                 | 0.1916  | 0.2219  | 0.3238  | 2194                 |
| \textit{Podocarpus hispaniolensis} | 5              | 0.1024  | 0.2529  | 0.3035  | 554                  |
|                          | 6               | 0.1716  | 0.0984  | 0.1799  | 336                  |
|                          | 9               | 0.2300  | 0.1150  | 0.2300  | 722                  |
|                          | 10              | 0.1822  | 0.3005  | 0.3874  | 3189                 |
|                          | 11              | 0.1301  | 0.2319  | 0.2999  | 475                  |
| Average                  |                 | 0.1633  | 0.1997  | 0.2855  | 1055                 |

Note: Collection site numbers as in Table 1 and Figure 1. Abbreviations: CC, Cordillera Central; CS, Cordillera Septentrional; $H_o$, expected heterozygosity; $H_e$, observed heterozygosity; SB, Sierra Bahoruco; SN, Sierra Neiba; $\pi$, nucleotide polymorphism diversity.
and $K = 5$ showed a clear separation between individuals from collection Sites 10 and 11 within Cordillera Septentrional, and those from Cordillera Central (Figure S1). For $K = 2$, individuals from Cordillera Central did not show any admixture but individuals from Cordillera Septentrional did, with one individual showing ca. 90% association with the Cordillera Central cluster. For $K = 5$, two individuals from Cordillera Central and five from Cordillera Septentrional showed admixture (Figure S1).

When both species and all loci were analysed, $K = 7$ received the highest support followed by $K = 5$. For both $K$ values, $P.\text{buchii}$ individuals from Sierra Bahoruco, Sierra Neiba and Cordillera Central belonged to different genetic clusters, but $P.\text{hispaniolensis}$ individuals from Cordillera Central and Cordillera Septentrional belonged to one genetic cluster (Figure 3a). $Podocarpus$ $\text{hispaniolensis}$ was nested within $P.\text{buchii}$ (Figure 4a,b). Both analyses recovered with high support ($PP = 1.0$ and $BS = 100\%$) a clade that included all $P.\text{hispaniolensis}$ sites.

### 3.3 Phylogenetic relationships

ML and BI analyses for the 11 collection sites recovered a congruent topology and showed that $P.\text{hispaniolensis}$ was nested within $P.\text{buchii}$ (Figure 4a,b). Both analyses recovered with high support ($PP = 1.0$ and $BS = 100\%$) a clade that included all $P.\text{hispaniolensis}$ sites.
Phylogenetic resolution and support were better for *P. hispaniolensis* sites than for *P. buchii* sites. The two southernmost *P. hispaniolensis* sites in Cordillera Central (5 and 6) were sister to a clade formed by the three northernmost *P. hispaniolensis* sites (9, 10 and 11). Both analyses also recovered Site 1 (*P. buchii* from Sierra Bahoruco Oriental) as sister to the *P. hispaniolensis* clade with high support (PP = 1.0 and BS = 99%). The relationships of *P. buchii* sites closer to the root showed some discrepancies between the two phylogenetic reconstruction analyses, but recovered both *P. buchii* Cordillera Central sites (7 and 8) as sister to each other with high support.

**FIGURE 3** Bar plots showing the genetic structure of both *Podocarpus* species on Hispaniola. Analyses using (a) all loci, (b) neutral loci and (c) loci under selection. Numbers at the top of the bar plots indicate collection sites.
(PP = 1.0 and BS = 95%), The ML and BI analyses for the five populations provided poor resolution and support due to extremely short branches (trees not shown).

3.4 | Demographic history of *Podocarpus* on Hispaniola

We discarded loci that had missing data in all individuals from a given population, as required by DIYABC, resulting in 244 SNPs for this analysis. The approximate Bayesian computation identified Models 6 and 8, the stepping-stone hypothesis with bottlenecks for each mountain colonization, as the best-fit under the direct and logistic regression approaches (Table S4 and Figure S2). Models that hypothesized an ancestral population from Cordillera Central that gave rise to two *Podocarpus* species or an allopatric origin for the two species in the two palaeo-islands reaching secondary contact in the Cordillera Central received no support (Figure 2, Table S4). The posterior predictive global error was 0.008 under the direct approach and 0.004 under the logistic regression approach. The prior predictive global error was 0.244 under the direct and logistic regression approaches. Direct approach type I errors were 0.06 and 0.2 for Models 6 and 8, respectively. Logistic type I errors were 0.036 and 0.1 for Models 6 and 8, respectively. Datasets from the posterior predictive distribution were close enough to the observed dataset for both Models 6 and 8 (Figures S3 and S4). Some summary statistics for both models had a proportion of simulated datasets below the observed one (Tables S5 and S6). The estimation of mean population sample sizes ranged from 386 to 1270 for Model 6 and 373 to 1840 for Model 8. Mean coalescent times were 15,500 years at t4, 1487 years at t3, 1010 years at t2 and 843 years at t1 for Model 6 (Table S7) and 3674 years at t4, 3325 years at t3, 3475 years at t2, and 1925 years at t1 for Model 8 (Table S8). Parameter estimates for Models 6 and 8 were accurate and unbiased, with low mean relative bias and RRMISE values (Tables S9 and S10, respectively).

4 | DISCUSSION

In the paleogeographically complex Caribbean island of Hispaniola, species and population divergence have been influenced by historical and current dispersal barriers. The genetic diversity of *Podocarpus* on Hispaniola was geographically structured by mountain ranges. This geographic structure, however, could not be attributed to Pliocene–Pleistocene marine incursions or the fusion of palaeo-islands because individuals did not form two genetic clusters or phylogenetic clades corresponding to the north and south palaeo-islands and because of the very recent estimated population divergence times. The genetic structure instead was likely influenced by contemporary topographic barriers of the dry valleys separating the different mountain ranges where *Podocarpus* grow. The demographic history of *Podocarpus* on Hispaniola was likely the product of a single colonization with subsequent dispersals to cloud forests in a stepping-stone manner with bottlenecks associated with each new mountain range migration. The phylogenetic reconstruction suggests a speciation scenario where *P. hispaniolensis* emerged from *P. buchii*.

4.1 | Patterns of genetic diversity and structure of *Podocarpus* on Hispaniola

The higher genetic diversity in *P. buchii* is likely due to its earlier diverging position with respect to the derived species *P. hispaniolensis* as shown in the phylogenetic tree of collection sites. A similar geographic distribution of genetic diversity was reported for Hispaniolan huitas, where southern populations had a higher nucleotide diversity (a) compared with northern populations (Brace et al., 2012). Yet, other taxa, such as the palm *Pseudophoenix vinifera* (Mart.) Becc., did not show significant differences in genetic diversity (*H*<sub>e</sub>) between northern and southern populations (Rodríguez-Peña et al., 2014).

Geographic distance is a major cause of genetic isolation (isolation by distance), which seems to occur in *Podocarpus* populations from Hispaniola. Even in small islands, such as Lord Howe Island in Australia, isolation by distance plays a role in lineage divergence in plants (Papadopulos et al., 2014). However, genetic distances are not always linearly correlated with geographic distances and other factors such as habitat configuration and maximum migration distance might influence isolation among populations (van Strien et al., 2015). This might explain why our most geographically distant collection sites (Sites 1 and 2 vs. 10 and 11) did not show the greatest genetic differentiation in our study. Overall genetic differentiation among collection sites (pairwise *F*<sub>ST</sub> values) was low.

Contrasting spatial patterns of genetic structure have been observed on Hispaniolan fauna. For example, in mammals of the genus *Plagiodonta* (Brace et al., 2012), and *Solenodon* (Turvey et al., 2016), and in birds (Sly et al., 2011), the distribution of genetically distinct allopatric lineages coincides with the geological fragmentation of the south and north palaeo-islands. Similarly, Matos-Maravi et al. (2014) showed that two vicariant events prompted the divergence between *Calisto* butterflies, one related to the uplift of the Cordillera Central, and a second one related to the repeated marine incursions in Neiba Valley-Cul de Sac Plain. In contrast, it is likely that for *Podocarpus* on Hispaniola, the dry valleys that lay in between the mountain chains acted as dispersal barriers that prompted divergence. This geographical structure is concordant with that of the lizard *Ameiva chrysoalaema* (Gifford et al., 2004), where the complex topography of Hispaniola formed three different evolutionary lineages.

The genetic admixture observed in *Podocarpus* suggests the presence of gene flow across the different mountain ranges or remnant ancestral similarity given the life cycle of *Podocarpus* and the fact that gymnosperms tend to have lower evolution rates (De La Torre et al., 2017). Other studies on Hispaniola have shown gene flow among bird populations occupying different mountain ranges, suggesting over-land dispersal between the south and north palaeo-islands (Sly et al., 2011). Despite the presence of dry valleys that might act as barriers for *Podocarpus*, gene flow might have
homogenized some populations; for example, the genetic cluster in blue (Figure 3a) is present in all populations. The different sets of loci used in the combined versus individual genetic structure analyses could explain the lack of admixture for *P. hispaniolensis* in the former but not the latter.

### 4.2 | Demographic history of *Podocarpus* on Hispaniola

The approximate Bayesian computation analysis supports a stepping-stone colonization model with bottleneck events with each new mountain range migration. However, the direction of colonization (south to north or vice versa) could not be elucidated unequivocally. A north to south colonization direction within Hispaniola was shown (south to north or vice versa) could not be elucidated unequivocally. However, the direction of colonization (e.g. Corsica and Crete, Steinbauer et al., 2013). Moreover, in other *Podocarpus* species it has been postulated that past climatic events may have been the cause of restrictions to gene flow between populations, which is now detected by the presence of unique haplодotic variants in nearby populations (Mellick et al., 2012; Quiroga et al., 2012) or distant populations (Dantas et al., 2015), even in divergent lineages since the Pleistocene (Migliore et al., 2020). Although we do not rule out the hypothesis that climatic events may have played a role in the divergence of these lineages, new studies with a historical/climatic focus should be conducted to analyse this hypothesis.

### 4.3 | Management units for conservation

*Podocarpus* on Hispaniola as other members of the genus, wrongly commonly named pines around the world, usually occur in mixed forests as small populations and isolated individuals that may result in different types of rarity and genetic polymorphism in conifers (Premoli et al., 2001). To the best of our knowledge, this is the first study using genotyping-by-sequencing of the whole genome within the genus *Podocarpus*, and SNP polymorphisms of endangered *P. buchii* and *P. hispaniolensis* yielded elevated within-population genomic diversity. This is despite the low number of individuals analysed per population that in some cases reflected the existence of solitary individuals in the field such as in collection Sites 2 and 9. Therefore, guidelines to protect them should focus at the individual and/or grove level. Moreover, not only Hispaniola *Podocarpus* had elevated heterozygosity but also all analysed populations had hundreds or in some cases thousands of private, that is unique, alleles. In contrast, reduced within-population genetic diversity was found in the only conifer found in north-eastern Brazil *Podocarpus sellowii* using SSR and ISSR markers (Dantas et al., 2015). Thus, highly variable genomic markers are useful in describing polymorphism levels even of populations occurring in small numbers. In addition, each pure *Podocarpus* taxon was clearly identified and among-population divergence, although relatively low, was clustered by location which can be used to design conservation actions. Similar geographically controlled genetic structure was documented in *Podocarpus parlatorei* from the subtropical Yungas of northern Argentina and southern Bolivia (Quiroga et al., 2012). Thus, conservation strategies for endangered *Podocarpus* should follow a multi-population approach where distinct population clusters of each species should be locally protected throughout Hispaniola. Southern populations of *P. buchii* at Sierra Bahoruco (south palaeo-island) and Sierra Neiba (south of north palaeo-island) although geographically close are genetically different and thus need to be conserved as distinct management units. In this area, only collection Site 3 at Loma del Quince is located into a National Park Sierra de Neiba (Figure S5). Similarly, nearly sympatric high-elevation *P. buchii* and low-elevation *P. hispaniolensis* from Cordillera Central, although they had lower number of unique alleles compared to allopatric southernmost *P. buchii* and northernmost *P. hispaniolensis*, were clearly identifiable by STRUCTURE and phylogenetic analyses. Therefore, Cordillera Central seems to reflect how the distinct species have differentiated by ecological divergence.
only the three collection sites of *P. hispaniolensis* in Cordillera Central are located in protected areas (National Park Montaña La Humeadora and in Las Neblinas Scientific Reserve; Figure S5). In addition, *P. buchii* from Barahona Province within Sierra Bahoruco Oriental and *P. hispaniolensis* from Yarao, Puerto Plata Province in Cordillera Septentrional have the greatest number of unique alleles (4908 and 3189, respectively) and thus can be considered reservoirs of neutral diversity, probably product of long-lasting isolation. Nonetheless, studies are needed to explore the adaptive value of such variants given that the relationship between a high neutral genetic diversity leading to a decreased risk of species extinction has been challenged (Teixeira & Huber, 2021). As more genetic information is provided by molecular markers at the entire genome levels, more difficult will be to use them for making conservation decisions given that different individuals will tend to be genetically unique and thus in need of conservation actions. Although our study comprises a restricted geographic area inhabited by two endangered conifers, conservation guidelines arise from the combination of genetic clustering methods, population uniqueness and degree of protection to better conserve their gene pools.

**ACKNOWLEDGEMENTS**

This work was funded by a Canadian NSERC-Discovery grant (RPIN-2014-03976) to JR, the School of Graduate Studies and the President’s Doctoral Investment fund of MUN to MENB; MPQ and ACP were supported by funds from the Agencia Nacional de Promoción Científica y Tecnológica de Argentina project PICT2019-149 and CONICET. We are grateful to the Royal Botanic Gardens Edinburgh, Royal Botanic Gardens Kew and the Montgomery Botanical Center for sharing leaf and DNA samples. We also want to thank Teodoro Clase for help in the field, and Francisco Jimenez from Jardín Botánico Nacional Dr. Rafael Ma. Moscoso for facilitating collections in the Dominican Republic. This research was facilitated by computer infrastructure from ACENET (www.ace-net.ca) and Compute Canada (www.computecanada.ca) for which Oliver Stueker provided support. This study is dedicated to the memory of Ricardo A. Nieto Sánchez. Open Access funding enabled and organized by Projekt DEAL.

**CONFLICT OF INTEREST**

The authors declare no conflict of interest.

**PEER REVIEW**

The peer review history for this article is available at https://publons.com/publon/10.1111/ddi.13444.

**DATA AVAILABILITY STATEMENT**

The data that support the findings of this study are available in Dryad at: https://doi.org/10.5061/dryad.37pmcvm3.

**ORCID**

Maria Esther Nieto-Blázquez [https://orcid.org/0000-0002-6201-282X](https://orcid.org/0000-0002-6201-282X)

**REFERENCES**

Abed, A., Légaré, G., Pomerleau, S., St-Cyr, J., Boyle, B., & Belzile, F. (2019). Genotyping-by-sequencing on the ion torrent platform in barley: methods and protocols. In W. Hardwood (Ed.), *Methods in molecular biology* (pp. 233–252). Humana Press.

Acevedo-Rodriguez, P., & Strong, M. T. (2008). Floristic richness and affinities in the West Indies. *Botanical Review*, 74, 5–36. https://doi.org/10.1007/s12229-008-9000-1

Blendinger, P. G. (2017). Functional equivalence in seed dispersal effective-ness of *Podocarpus parlatorei* in Andean fruit-eating bird assemblages. *Frontiers in Ecology and Evolution*, 5, 57. https://doi.org/10.3389/fevo.2017.00057

Brace, S., Barnes, I., Powell, A., Pearson, R., Woolwer, L. C., Thomas, M. G., & Turvey, S. T. (2012). Population history of the Hispaniolan hutia *Plagiodontia aedium* (Rodentia: Capromyidae): Testing the model of ancient differentiation on a geotectonically complex Caribbean island. *Molecular Ecology*, 21, 2239–2253.

Brown, J. L., Cameron, A., Yoder, A. D., & Vences, M. (2014). A necessarily complex model to explain the biogeography of the amphibians and reptiles of Madagascar. *Nature Communications*, 5, 5046. https://doi.org/10.1038/ncomms6046

Cano-Carmona, E., Veloz-Ramírez, A., & Cano-Ortiz, A. (2010). Contribution to the biogeography of the Hispaniola (Dominican Republic, Haiti). *Acta Botánica Gallica*, 157, 581–598. https://doi.org/10.1080/12538078.2010.10516233

Cano-Ortiz, A., Musarella, C. M., Fuentes, J. C. P., Gomes, C. J. P., & Cano, E. (2016). Distribution patterns of endemic flora to define hotspots on Hispaniola. *Systematics and Biodiversity*, 14, 261–275. https://doi.org/10.1080/14772000.2015.1135195

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22, 3124–3140. https://doi.org/10.1111/mec.12354

Chessell, D., Dufour, A. B., & Thioulouesse, J. (2004). The ade4 package - I: One-table methods. *R News*, 4, 5–10.

Cornuet, J.-M., Dudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Lebois, R., Marin, J.-M., & Estoup, A. (2014). DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, 30, 1187–1189. https://doi.org/10.1093/bioinformatics/btt763

Dantas, L. G., Esposto, T., de Sousa, A. C. B., Félix, L., Amorim, L. L. B., Benko-Iseppon, A. M., Batalha-Filho, H., & Pedrosa-Harand, A. (2015). Low genetic diversity and high differentiation among relic populations of the neotropical gymnosperm *Podocarpus sellowii* (Klotz.) in the Atlantic Forest. *Genetica*, 143, 21–30. https://doi.org/10.1007/s10709-014-9809-y

De La Torre, A. R., Li, Z., Van de Peer, Y., & Ingvarsson, P. K. (2014). A necessarily complex model to explain the biogeography of the Hispaniola (Dominican Republic, Haiti). *Acta Botánica Gallica*, 157, 581–598. https://doi.org/10.1080/12538078.2010.10516233

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361. https://doi.org/10.1007/s12229-011-9548-7

Enright, N. J., & Jaffré, T. (2011). Ecology and distribution of the Malesian Podocarps. In B. L. Turner, & L. A. Cernusak (Eds.), *Podocarpaceae in tropical forests* (pp. 57–77). Smithsonian Institution Scholarly Press.

Ersts, P. J. (2011). *Geographic distance matrix generator v1.2.3*. American Museum of Natural History, Center for Biodiversity and Conservation. http://biodiversityinformatics.amnh.org/open_source/gdmg

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A
Patterns and processes. Fascinating life sciences (pp. 521–546). Springer.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology, 61, 539–542. https://doi.org/10.1093/sysbio/sys029

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. Molecular Ecology Notes, 4, 137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Russell, F. A. (1991). Land use: Erosion and poverty in the Dominican Republic. The Centennial Review, 35, 323–334.

Sly, N. D., Townsend, A. K., Rimmer, C. C., Jason, M., Latta, S. C., & Lovette, I. J. (2011). Ancient islands and modern invasions: Disparate phylogeographic histories among Hispaniola’s endemic birds. Molecular Ecology, 20, 5012–5024. https://doi.org/10.1111/j.1365-294X.2011.05073.x

Sly, N. D., Townsend, A. K., Rimmer, C. C., Townsend, J. M., Latta, S. C., & Lovette, I. J. (2010). Phylogeography and conservation of the endemic Hispaniolan Palm-Tanagers (Aves: Phaenicoptilus). Conservation Genetics, 11, 2121–2129. https://doi.org/10.1007/s10592-010-0098-4

Spielman, D., Brook, B. W., & Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. Proceedings of the National Academy of Sciences of the United States of America, 101, 15261–15264.

Stamatakis, A. (2006). RAxML VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics, 22, 2688–2690. https://doi.org/10.1093/bioinformatics/btl446

Steinbauer, M. J., Irl, S. D. H., & Beierkuhnlein, C. (2013). Elevation-driven ecological isolation promotes diversification on Mediterranean islands. Acta Oecologica, 47, 52–56. https://doi.org/10.1016/j.actao.2012.11.004

Supple, M. A., & Shapiro, B. (2018). Conservation of biodiversity in the genomics era. Genome Biology, 19, 131. https://doi.org/10.1186/s13059-018-1520-3

Teixeira, J. C., & Huber, C. D. (2021). The inflated significance of neutral genetic diversity in conservation genetics. Proceedings of the National Academy of Sciences of the United States of America, 118, e2015096118.

Townsend, A. K., Rimmer, C. C., Latta, S. C., & Lovette, I. J. (2007). Ancient differentiation in the single-island avian radiation of endemic Hispaniolan chat-tanagers (Aves: Calyptophilus). Molecular Ecology, 16, 3634–3642.

Turvey, S. T., Peters, S., Brace, S., Young, R. P., Crumpton, N., Hansford, J., Nuñez-Miño, J. M., King, G., Tsalikidis, K., Ottenwalder, J. A., Timpson, A., Funk, S. M., Brocca, J. L., Thomas, M. G., & Barnes, I. (2016). Independent evolutionary histories in allopatric populations of a threatened Caribbean land mammal. Diversity and Distributions, 22, 589–602. https://doi.org/10.1111/ddi.12420.

van Strien, M. J., Holderegger, R., & van Heck, H. J. (2015). Isolation-by-distance in landscapes: Considerations for landscape genetics. Heredity, 114, 27–37. https://doi.org/10.1038/hdy.2014.62

BIOSKETCH
Maria Esther Nieto-Blázquez is an evolutionary biologist broadly interested in biogeography, systematics and genomics. This work is part of her doctoral studies at Memorial University of Newfoundland. She is currently a postdoc at Senckenberg Biodiversity and Climate Research Centre, Frankfurt am Main, working on population genomics of wildcat (Felis silvestris) and beech (Fagus sylvatica).

SUPPORTING INFORMATION
Additional supporting information may be found in the online version of the article at the publisher’s website.

How to cite this article: Nieto-Blázquez, M. E., Quiroga, M. P., Premoli, A. C., & Roncal, J. (2021). Podocarpus in the palaeogeographically complex island of Hispaniola: A stepping-stone colonization and conservation recommendations. Diversity and Distributions, 00, 1–13. https://doi.org/10.1111/ddi.13444