Plant Ecology and Evolution 151 (1): 61–76, 2018
https://doi.org/10.5091/plecevo.2018.1343

PLANT ECOLOGY EVOLUTION

REGULAR PAPER

Genetic variation and dispersal patterns in three varieties of *Pinus caribaea* (Pinaceae) in the Caribbean Basin

Virginia Rebolledo Camacho[1,3], Lev Jardón Barbolla[2], Ivón Ramírez Morillo[3], Alejandra Vázquez-Lobo[4], Daniel Piñero[5] & Patricia Delgado[6,*]

[1]Instituto de Investigaciones Forestales, Universidad Veracruzana, Parque Ecológico "El Haya", Carretera antigua a Coatepec S/N, CP 91000 Xalapa, Veracruz, México
[2]Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades, Universidad Nacional Autónoma de México, Torre II de Humanidades, 4º piso, Ciudad Universitaria, CP 04510 Ciudad de México, México
[3]Centro de Investigación Científica de Yucatán, A.C. Calle 43 #130 x 32 y 34, Colonia Chuburná de Hidalgo, CP 97205 Mérida, Yucatán, México
[4]Centro de Investigación en Biodiversidad y Conservación, Universidad Autónoma del Estado de Morelos, Av. Universidad No. 1001, Colonia Chamilpa, CP 62209 Cuernavaca, Morelos, México
[5]Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Tercer Circuito Exterior S/N, Ciudad Universitaria, CP 04510 Ciudad de México, México
[6]Facultad de Agrobiología "Presidente Juárez", Universidad Michoacana de San Nicolás de Hidalgo, Paseo de la Revolución esquina con Berlín S/N, Colonia Viveros, CP 60170 Uruapan, Michoacán, México
*Author for correspondence: dvalerio@umich.mx

Background – *Pinus caribaea* Morelet comprises three varieties of tropical pines distributed in the Caribbean Basin: *P. caribaea* var. *hondurensis*, var. *caribaea*, and var. *bahamensis*. The insular and continental distribution of these varieties, as well as the geological processes in the region, have been important factors for analysing evolutionary processes implicated in the diversification of these lineages. In this study, we evaluate the genetic and geographic structure within and between these three varieties in order to infer the possible origin and dispersal routes of these taxa.
Methods – We used six polymorphic nuclear microsatellites (nSSR) in fifteen representative populations of the three pine varieties, sampled throughout their natural range in Central America, Cuba and the Bahamas islands.
Results – The varieties contain similar levels of genetic variation (mean H_e = 0.571), with several populations out of Hardy-Weinberg equilibrium, and significant levels of inbreeding (0.097–0.184, $P \leq 0.05$). A slight but significant genetic differentiation was found between the varieties (R_{ST} = 0.088) and populations (R_{ST} = 0.082), and genetic differentiation increased with geographic distance (r^2 = 0.263). Distance and Bayesian BAPS analyses generated seven groups; two represented by the two island varieties and the remainder by the Central American populations of var. *hondurensis*. Migration rate estimates between pairs of groups ranged from M = 0.47 to M = 20.16. Estimates were generally higher from the continent to islands, with the highest migration rate estimated from a continental genetic group to the Cuba island group of var. *hondurensis* (M = 20.16).
Conclusions – This study supports the hypothesis of a recent origin of these pine taxa through the migration of an ancestor from Central America, where the historical demography is associated with events of colonization, expansion and contraction of populations. The genetic variation and differentiation suggest that the three varieties are divergent lineages that currently share allelic variants, indicating that their speciation has not yet completed.

Key words – *Pinus caribaea* varieties, genetic variation, microsatellites, lineage divergence, migration routes, Caribbean Basin.

ISSN: 2032-3913 (print) – 2032-3921 (online)

INTRODUCTION

Mexico is considered a secondary centre of diversification of the genus *Pinus* L., represented by 52 species, two subspecies, 14 varieties and four forms (Perry 1991, Gernandt & Pérez de la Rosa 2014). This high diversity originated during the early Oligocene (≈37 Mya) due to multiple migrations of species from the North American temperate zone to Mexico, with subsequent local speciation events and adaptations to different climatic and geological conditions (Millar 1993, Farjon 1996, Dvorak et al. 2009), including species capable of establishing at the lowest altitude (sea level) (Farjon 2005). The species *Pinus caribaea* Morelet is naturally distributed at altitudes from sea-level to 700 metres above sea-level (m a.s.l) (rarely reaching 1000 m a.s.l) in tropical forests susceptible to seasonal flooding (lowland and coastal populations) and on a variety of soil types. Soils range from sandy, acidic, nutrient-poor to deep sandy clays with good drainage and to pure clays (Van 2002, Farjon & Styles 1997). *Pinus caribaea* is subdivided into three varieties with a wide distribution in the Caribbean Basin: *Pinus caribaea* var. *hondurensis* (Sénécl.) W.H.Barret & Golfari (Mexico and Central America), *Pinus caribaea* var. *caribaea* (Western Cuba and Juventud Island), and *Pinus caribaea* var. *bahamensis* (Griseb.) W.H.Barret (Bahamas, Turks and Caicos Islands) (Nikles 1966, Farjon & Styles 1997). Taxonomically, *P. caribaea* and its varieties belong to *Pinus* subsection *Australes*, together with *P. occidentalis* Swartz, *P. maestrensis* Bisse, *P. cubensis* Griseb, *P. elliottii* Engelm, *P. echinata* Mill., *P. palustris* Mill., *P. pungens* Lamb., *P. taeda* L., *P. oocarpa* Schiede ex Schltdl. and *P. tecunumani* F. Schwerdtf. ex Eguiluz & J.P.Perry (Dvorak et al. 2000a, Gernandt et al. 2005, Hernández-León et al. 2013). Most of the populations are monospecific, but sometimes *P. caribaea* var. *hondurensis* is in sympatry with *P. oocarpa* or *P. tecunumani* (Honduras, Belize, Guatemala, El Salvador and Nicaragua) (Perry 1991, Farjon 1996, Dvorak et al. 2000b). *Pinus caribaea* var. *hondurensis* is the taxon with the most southern distribution in America (Nicaragua) and, while this variety has a wider distribution and its conservation status is considered as least concern (LC, IUCN 2016), present logging activities threaten to fragment and reduce populations in some areas (Farjon 2013). *Pinus caribaea* var. *caribaea* and *P. caribaea* var. *bahamensis* are threatened throughout their natural distribution; *P. caribaea* var. *caribaea* is considered as endangered (EN) because of logging and conversion to pasture, especially on the mainland of Cuba, while *P. caribaea* var. *bahamensis* is now considered as vulnerable (VU) and could be at risk due to population decline in the Turks and Caicos Islands as result of attack by the pine tortoise scale, *Toumeyella parvicornis* (Cockerell, 1897) (Farjon 2013, Sánchez et al. 2014, IUCN 2016).

Morphologically, the three varieties are very similar, with 2–5 needles per fascicle, but they differ mainly by a seedling stage phase, the number and position of resin canals on the leaves and certain seed wing features (Farjon & Styles 1997). Some authors have found molecular differences among the three varieties, which has an evolutionary implication. Using isoenzyme markers, Zheng & Ennos (1999) analysed variation and genetic relationships in populations of the varieties *caribaea* and *bahamensis*, while Matheson et al. (1989) examined populations of the varieties *hondurensis* and *bahamensis*. These studies reported significant differences in the presence and frequency of alleles between the varieties, supporting the notion that the varieties are different evolutionary entities or units. When the three varieties of *P. caribaea* and three additional species of pines distributed in the Caribbean Basin (*P. cubensis*, *P. maestrensis* and *P. occidentalis*) were analysed using a phylogeographic approach with plastid microsatellites, the highest genetic diversity was observed in the populations of *P. caribaea* var. *hondurensis* distributed in Central America (H_e=0.951) and different degrees of genetic differentiation were found among the varieties and species (R_{ST} = 0.230 among varieties, R_{ST} = 0.110 among species) (Jardón-Barbolla et al. 2011).

This differentiation is probably associated with geological and climatic processes of the Caribbean Basin that have acted either as barriers or as corridors of dispersion. While the geological history of this region has been widely studied, nonetheless, it remains poorly understood. It has been proposed that the Greater Antilles and the Bahamas have been separated from North and Central America since the Miocene (≈23 Mya) (Iturralde-Vinent 2006, Pindell et al. 2006). However, paleogeographic studies show that changes in sea level during recent glacial periods could have left extensive areas of emerged land with temporary connections (coral reef land bridges) between the islands and Central America (Schuchert 1935, Hedges 1996a, 1996b). This theory, along with the distribution of the related species of *Pinus* subsection *Australes* in Central America, suggests that the origin and distribution of contemporary biota in this region is a result of species migration events from the continent to the islands at a geological time of less than 20 Mya (Mirov 1967, Nikles 1966, Hedges 1996b, Iturralde-Vinent 2004–2005, Moonlight et al. 2015). Considering that *Pinus* arrived to Mexico and Central America between 35 and 20 Mya (Millar 1993, Hedges 1996b), we can assume that *Pinus caribaea* varieties, and subsection *Australes* in general, are of relatively recent origin.

To date, two hypotheses have been proposed concerning the origin of the species classified in *Pinus* subsection *Australes*. The first is based on a phylogenetic reconstruction obtained from morphological data (Adams & Jackson 1997), which proposed that the ancestor originated in the southeastern region of North America, migrated from southern Florida to the Caribbean islands and then to Central America. The second hypothesis, initially suggested by Nikles (1966) and Mirov (1967), was based on nuclear dominant RAPD (Random Amplified Polymorphic DNA) markers (Dvorak et al. 2000a) and chloroplast DNA microsatellite data (Jardón-Barbolla et al. 2011), and proposed a Central American origin of the ancestor: the ancestral *P. caribaea* would have been separated prior to the divergence of Australes, then migrated to Central America through the Western Gulf of Mexico. This early form is suggested to have migrated to the Caribbean and eventually to the Florida peninsula, from where the ancestor of *P. caribaea* var. *caribaea* would have dispersed (probably infrequently over an extended period after the Pleistocene) to Cuba and the Bahamas (Dvorak et al. 2000a). It has also been inferred that populations of *caribaea* and *ba-*

hamensis varieties have had recent colonization events to the islands of Cuba and the Bahamas (97 900 years BP), whereas *P. caribaea* var. *hondurensis* populations are older, associated with expansion processes during lower temperatures at the beginning of a glacial period (326 300 years BP; Jardón-Barbolla et al. 2011). Interpretations of these data should consider that chloroplast markers are haploid and inherited parentally through pollen, thus their coalescence time inferences are shorter than those estimated with diploid nuclear DNA (Rosenberg & Nordborg 2002, Avise 2009). It is therefore relevant to study the nuclear genetic variation in order to achieve a more comprehensive historical reconstruction of evolutionary processes. For example, in a phylogeographic study of *Pinus strobus* L. (Zinck & Rajora 2016), inferences with both nuclear and chloroplast markers were consistent, revealing a single refuge, two re-colonization routes and three genetically distinguishable lineages. In contrast to the chloroplast markers, nuclear markers allowed the detection of higher genetic diversity and more pronounced levels of genetic structure. Similarly, joint use of mitochondrial DNA sequences and nuclear loci allowed the detection of historical introgression events in *Picea obovata* Ledeb. and *Picea abies* (L.) H.Karst. that were missed when using maternally inherited markers alone (Tsuda et al. 2016).

In this sense, nuclear codominant microsatellites (nuclear simple sequence repeats, nSSR) are suitable for providing complementary information with which contrast the different hypotheses about the origin of *P. caribaea*, considering their high mutation rates from 10^{-5} to 10^{-3} mutations per site per generation (Schlötterer 2000, Boys et al. 2005). Nuclear SSRs are selectively neutral, conserved across species and have been widely used to evaluate genetic variation and structure, gene flow, inbreeding and effective population sizes (Karhu 2001, Boys et al. 2005, Furlan et al. 2007). Moreover, nSSRs allow the inference of demographic processes such as dispersal, migration, expansion, or fragmentation of populations (Adams & Jackson 1997, Rajora et al. 2000, Williams et al. 2000, Mariette et al. 2001, Shepherd et al. 2002, Al-Rabab'ah 2003, Boys et al. 2005, Karhu et al. 2006, Dvorak et al. 2009, Wang et al. 2009, Sánchez et al. 2014, Zinck & Rajora 2016, Tsuda et al. 2016). Currently, only three studies have been conducted on patterns of genetic variation within the varieties *hondurensis* and *bahamensis* with nSSRs. In three experimental *P. caribaea* var. *hondurensis* populations established in Brazil with individuals from Poptún, Guatemala, very low levels of variation and genetic differentiation were detected (H_e= 0.249, F_{ST} = 0.021; Furlan et al. 2007). Likewise, low levels of genetic variation and differentiation were found for two populations of the same variety distributed in Mexico (H_e = 0.465, R_{ST} =0.033) and this was attributed to a recent reduction of effective population sizes (8100 to 35 000 years ago; Delgado et al. 2011). On the other hand, var. *bahamensis* showed high levels of genetic variation and structure among populations from the Bahamas, and Turks and Caicos islands (TCI), which was attributed to the effect of geographic isolation of the population distributed in the latter region (Sánchez et al. 2014). The joint study of the three varieties is necessary both for a better understanding of the evolutionary history of this taxon, as well as to outline management and conservation schemes of this important complex of Caribbean pines.

In this study, nuclear microsatellites were used to test the hypothesis that the three varieties of *P. caribaea* (var. *caribaea*, *hondurensis* and *bahamensis*) distributed in the Caribbean Basin, represent independent evolutionary lineages originating from one ancestor distributed in Central America. Our aims were (i) to estimate levels of genetic variation in the varieties and populations, inbreeding indices and effective population sizes, and (ii) to determine the geographic structure of genetic variation across populations to infer possible migration routes (gene flow). Finally, we analysed and discussed the geographic and demographic processes that may underlie the distribution of genetic variation in this group of pines.

MATERIALS AND METHODS

Study populations and sampling

Fifteen populations of the three varieties of *Pinus caribaea* were sampled and geo-referenced throughout the Caribbean Basin (fig. 1). Leaf material was collected from two populations of *P. caribaea* var. *caribaea* located at the western end of the island of Cuba (Viñales and Mil Cumbres), from two populations of *P. caribaea* var. *bahamensis* in the Bahamas (islands of Andros and New Providence), and from eleven populations of *P. caribaea* var. *hondurensis*, one in Mexico and one in Guatemala, four in Belize, three in Honduras, and two in Nicaragua. Of these, seven were located inland (H1, H2, H4, H6, H8, H9 and H10) and the other four were coastal lowland sources (table 1). Needles were collected from 13 to 30 trees in each of the population respecting a minimum distance of 50 m between trees, in order to reduce the probability of parentage (Flores et al. 2005). A total of 316 individuals were used in the study. Plant tissue was stored in plastic bags at -80°C for subsequent DNA extraction.

DNA extraction, amplification and genotyping

Genomic DNA was extracted with a CTAB miniprep method (Vázquez-Lobo 1996). The seven nSSRs assayed were derived from *Pinus taeda* (Elsik et al. 2000); PtTX3025, PtTX3013, PtTX3020, PtTX2146, PtTX2123, PtTX3029 and PtTX2037. The PCR amplification reaction conditions described below were performed using a MasterCycler Gradient thermocycler (Eppendorf Inc), according to Elsik et al. (2000), with modifications in the concentration of magnesium chloride (4 mM). Specific touchdown PCR conditions were as follows: one cycle at 94°C for 5 min; two cycles of 94°C for 1 min, 60°C for 1 min and 72°C for 35 s; 20 cycles of 94°C for 45 s, 45 s at specific annealing temperature (TM) of the primer pair, decreasing by 0.5°C each cycle, and 72°C for 1 min; 20 cycles at 94°C for 1 min, final TM for 1 min and 72°C for 1 min; final extension at 72°C for 5 min. The annealing temperatures for each primer pair were; 63°C for PtTX3013, PtTX3029, PtTX2037, 59°C for PtTX3025, 57°C for PtTX2146, PtTX2123, and 65°C for PtTX3020. The fragments were separated by electrophoresis on 5% polyacrylamide gels (7M Urea; Tris-Borate-EDTA [TBE] buffer at 0.5%), and run at 50–60 W for 1.5–3.5 hours, depending

on fragment size. A positive control (genotype) was used for each nSSR to confirm and standardize the allele sizes. Fragments were revealed with the silver nitrate staining method (Echt et al. 1996), and fragment size was determined visually using a 10 bp DNA ladder (Invitrogen).

The presence of null alleles (non-amplified alleles) has been reported in two of seven microsatellites used here (PtTX2037 and PtTX3020; Williams et al. 2000, Shepherd et al. 2002); therefore, the frequency of null alleles for all of the loci was estimated using Micro-Checker v. 2.2.3 and the genotypes adjusted according to the correction algorithm of Brookfield (van Oosterhout et al. 2004). A low proportion of null alleles was determined for the loci PtTX2146 (0.171) and PtTX2037 (0.204) and a high proportion for PtTX3029 (0.743). According to studies in which null alleles were present, it is possible to correct their effect (detecting a significant heterozygosity deficit relative to Hardy-Weinberg equilibrium, which could be misconstrued as evidence of inbreeding) by eliminating some individuals with high proportions of missing data, and repeating the analysis (Williams et al. 2000, Dakin & Avise 2004). Therefore, 17 individuals were excluded for all loci; null alleles for the locus PtTX2146 were corrected in five populations (H5, H6, H8, H10 and H11), and for the locus PtTX2037 in four populations (H1, H10, C2 and B2). However, it was not possible to correct this problem in locus PtTX3029 in eight populations (H2, H4, H5, H6, H7, H8, C1, C2), which maintained high proportions of null alleles (0.214 to 0.373), and this locus was therefore eliminated from the study altogether. As a result of this correction, all of the analyses of variation and genetic structure were carried out with six loci and a total of 299 individuals.

Data analyses

Genetic variation was estimated using the following parameters: number of alleles per locus (A), average number of alleles per locus (n), number of effective alleles per locus (A_e), observed (H_o) and expected (H_e) heterozygosity. The inbreeding index (F_{IS}) was estimated according to Wright (1965), and deviations from Hardy-Weinberg equilibrium were assessed with an unbiased estimation using the Markov chain Monte Carlo (MCMC) method with 100 000 steps (Guo & Thompson 1992). These analyses were performed

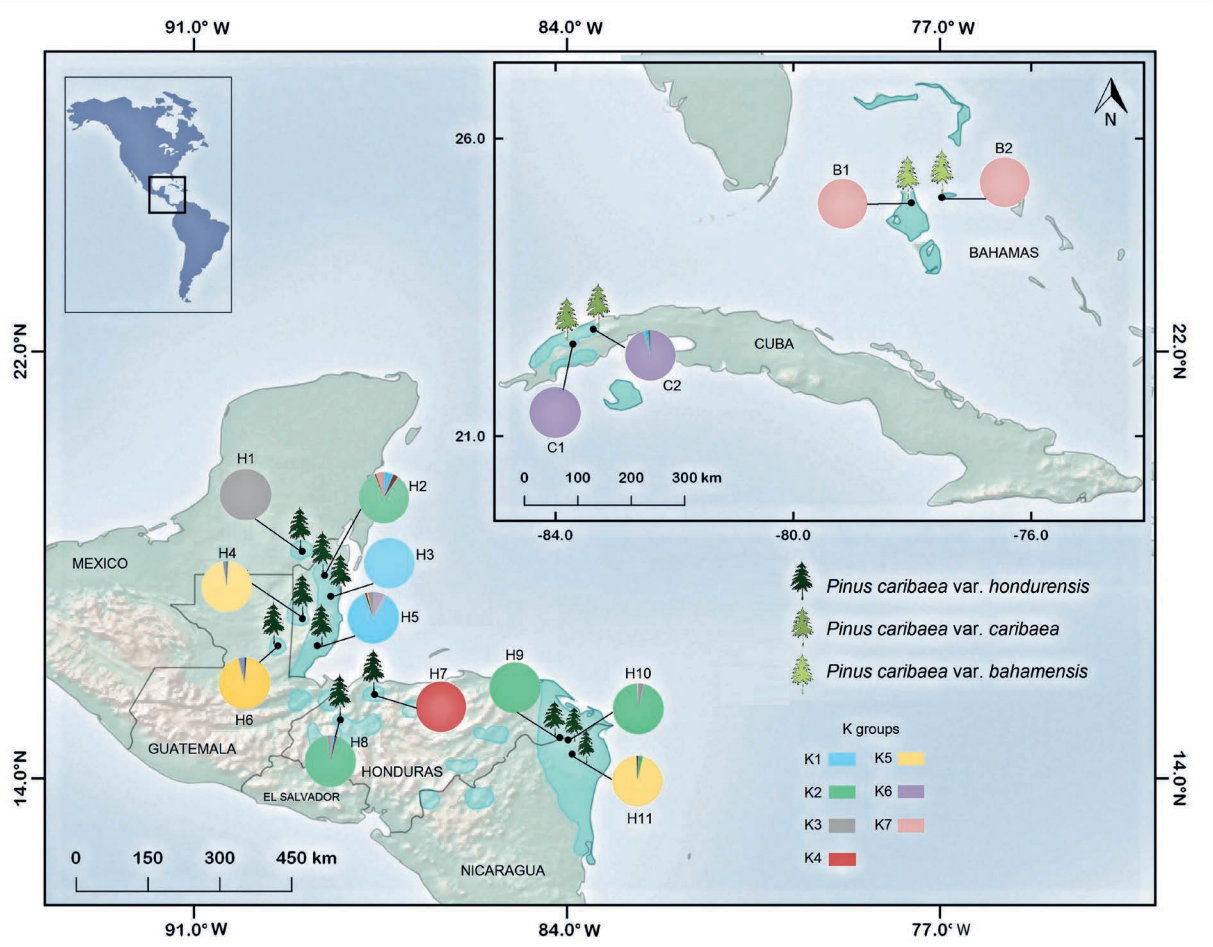

Figure 1 – Geographical location of fifteen populations for *Pinus caribaea* varieties distributed in the Caribbean Basin. The graphics represent the proportion of individuals for each population in accord to the best clustering (K seven groups) obtained with BAPS analysis (*log*ML = -3612; P = 1.000). Population codes according to table 1. Blue colour on the map represents the geographic distribution range of the varieties (adapted from Francis 1992). Date and projection of the map: WGS 1984, Latitude/Longitude. Map created with QGIS version 2.14 (QGIS Development Team 2017).

Table 1 – Geographical location and genetic parameters estimated in fifteen populations of the three *Pinus caribaea* varieties distributed in the Caribbean Basin.

Alt, altitude; N, sample size; *n*, total number of alleles; *A*, average number of alleles; A_r, allelic richness; A_e, average number of effective alleles per locus; H_o and H_e, average of heterozygosis observed and expected; F_{IS}, inbreeding index, * *P*= 0.05, ***P*= 0.01. SD, standard deviations are given in parentheses and CI-95%, confidence intervals in brackets.

Population	Country/ State, District or Department	Latitude Longitude	Altitude (m s.a.l)	*N*	*n*	*A*	A_r	A_e	H_o	H_e	F_{IS}
Pinus caribaea var. *hondurensis*											
H1-Caobas	Mexico/ Quintana Roo	18º14′77.5″N 88º57′60.4″W	35	17	18	3	2.5	2.1	0.421 (0.264)	0.488 (0.219)	0.057 [0.001–0.221]
H2-Carmelitas	Belize/Belize	17°48′12.2″N 88°32′55.0″W	13	19	19	3.2	2.7	2.4	0.441 (0.252)	0.568 (0.132)	0.233** [0.006–0.282]
H3-Rock	Belize/Belize	17°24′43.1″N 88°26′02.4″W	16	18	21	3.5	2.8	2.6	0.523 (0.182)	0.593 (0.147)	0.124 [0.013–0.253]
H4-Mountain Pine	Belize/Cayo	16°59′35.0″N 88°57′50.1″W	501	30	24	4	2.8	2.6	0.566 (0.106)	0.581 (0.138)	0.028 [0.000–0.125]
H5-Deep River	Belize/Toledo	16°29′16.0″N 88°41′08.3″W	31	22	23	3.8	2.9	2.7	0.491 (0.238)	0.578 (0.246)	0.154* [0.051–0.289]
H6-Dolores	Guatemala/ El Petén	16°29′07.7″N 89°25′34.9″W	438	20	25	4.2	3.2	2.9	0.537 (0.111)	0.634 (0.140)	0.157* [0.004–0.253]
H7-Mezapa	Honduras/ Atlántida	15°34′18.6″N 87°36′38.2″W	306	25	24	4	3.0	2.9	0.521 (0.254)	0.600 (0.238)	0.135* [0.020–0.245]
H8-Trinidad	Honduras/ Santa Bárbara	15°05′54.0″N 88°15′10.0″W	200	16	22	3.7	2.9	2.6	0.453 (0.188)	0.589 (0.172)	0.239* [0.078–0.352]
H9-Leimus	Honduras/ Gracias a Dios	14°45′57.0″N 84°08′08.7″W	90	21	22	3.7	2.7	2.3	0.516 (0.157)	0.553 (0.141)	0.068 [0.000–0.146]
H10-Waspam	Nicaragua/ North Atlantic	14°43′09.8″N 83°58′47.1″W	87	24	22	3.7	2.9	2.6	0.537 (0.124)	0.621 (0.073)	0.140* [0.007–0.205]
H11-Moss	Nicaragua/ North Atlantic	14º27′13.8″N 83°54′14.4″W	128	21	23	3.8	3.0	2.5	0.516 (0.153)	0.616 (0.070)	0.177* [0.005–0.246]
Average					22	3.7	2.8	2.6	0.502	0.575	0.124** [0.034–0.225]
Pinus caribaea var. *caribaea*											
C1-Viñales	Cuba/ Pinar del Río	22°32′48.4″N 83°42′29.5″W	239	19	20	3.3	2.7	2.4	0.522 (0.194)	0.573 (0.123)	0.092 [0.000–0.126]
C2-Mil Cumbres	Cuba/ Pinar del Río	22°47′45.9″N 83°21′57.4″W	185	18	19	3.2	2.7	2.3	0.384 (0.161)	0.555 (0.173)	0.319** [0.246–0.518]
Average					19.5	3.2	2.7	2.3	0.476	0.564	0.184* [0.025–0.349]
Pinus caribaea var. bahamensis											
B1-Andros	Bahamas	25°00′31.2″N 77°30′06.9″W	9	16	19	3.2	2.6	2.3	0.591 (0.145)	0.551 (0.112)	-0.063 [0.000–0.112]
B2-New Providence	Bahamas	24°55′13.1″N 78°00′49.8″W	4	13	19	3.2	2.7	2.3	0.433 (0.195)	0.593 (0.104)	0.280** [0.017–0.368]
Average					19	3.2	2.7	2.3	0.512	0.576.	0.097 [-0.079–0.201]

using Arlequin v. 3.5.1.2 (Excoffier & Lischer 2010). Further, ADZ v.1 (Szpiech et al. 2008) was used to estimate allelic richness, A_r estimation, using a rarefaction approach to standardize estimates to the smallest population sample size of the data set (El Mousadik & Petit 1996).

Genetic effects of population demographic decline were examined using the T_2 statistic (Cornuet & Luikart 1996), which reflects the deviation from expectations at demographic equilibrium (Budde et al. 2017). The test was performed using the infinite allele model (IAM), the stepwise mutation model (SMM) and the two-phase model (TPM; 70% of mutations under the SMM model and 30% under IAM) with Bottleneck v.1.2.02 (Piry et al. 1999). Significance in the three mutation models was tested using Wilcoxon´s signed rank test, with 10 000 replicates.

Genetic structure (R_{ST}), was estimated with a hierarchical analysis of molecular variance (AMOVA) assuming a stepwise mutation model (SMM; Slatkin 1995). The analysis was divided into four components: among the three varieties (F_{CT}), between populations within varieties (F_{SC}), between

individuals within populations of each variety (F_{IS}) and within individuals (F_{IT}). This analysis was also conducted between all populations and between the *K* groups obtained with BAPS analysis (see next paragraph), assessing three components; between populations or *K* Groups (F_{CT}), between individuals within populations or *K* groups (F_{IS}) and within individuals (F_{IT}). All statistical significance was obtained with 1000 non-parametric permutations (Excoffier & Lischer 2010).

The association between geographic structure and genetic variation of populations and varieties was estimated with the Bayesian inference algorithm implemented in the program BAPS v. 5.4 (Corander et al. 2008). The algorithm defines groups of populations using information pertaining to their spatial distribution in order to detect the most likely genetic structure. The estimates were obtained with a spatial clustering method, assuming 1 to 15 groups (*K*), with 10 replicates per *K*, using 10 000 iterations for estimates, preceded by 1000 iterations discarded as burn-in. The partition with the highest marginal probability (*Log*ML; natural logarithm of likelihood) was selected as the one that best describes the genetic structure of the data. In order to analyse the genetic distance between populations, a distance tree using the neighbor-joining method was constructed with the POPTREE2 software (Takezaki et al. 2010), based on standardized genetic distances (*Da*) (Nei et al. 1983). Support of this distance tree was evaluated by bootstrap analysis, using 1000 replicates (Takezaki & Nei 1996).

Effective population size (N_e) and historical migration rates (*M*) were estimated for the groups of populations obtained with BAPS. Estimations were carried out with Migrate version 3.4.2 (Beerli 2008, software accessed in January 2013), under a Bayesian approach (Beerli & Palczewski 2010). Uniform priors were used for all parameters with three independent runs to verify the convergence. Markov chains were obtained with 500 000 iterations, after a burn-in period of 10 000 steps, and a thinning interval of 0.0 to 100 (Beerli 2008). A mutation rate (μ) of 10^{-3} per generation was assumed; this rate has been used with nSSRs in other pine species (Boys et al. 2005, Delgado et al. 2011). Theta (θ) and *M* parameters were generated with the F_{ST} calculation ($F_{ST} = 1/1 + 4N_em$; Beerli 1998, 2008), and since $\theta = 4N_e\mu$, N_e was estimated as $\theta/4 \times 10^{-3}$ (Boys et al. 2005). In addition, an analysis of isolation by distance (IBD) was performed, regressing the genetic distance between pairs of populations on their geographic distance and testing the relationships using a Mantel test, with 10 000 permutations (Mantel 1967, Sokal & Rohlf 1995). Standardized genetic distances (*Da*) were used for the genetic data (Nei et al. 1983), and absolute distances (in kilometres) through Mercator transformation were used for geographic distances (ESRI 1992–2000). This analysis was run with the IBD program (Bohonak 2002).

RESULTS

Genetic variation

The six loci analysed were polymorphic; four had high levels of genetic diversity (expected heterozygosity): PtTX2146 (H_e = 0.732), PtTX2123 (H_e = 0.679), PtTX3020 (H_e = 0.659) and PtTX3025 (H_e = 0.564). All loci presented genetic diversity values of $H_e > 0.5$. A total of $n = 35$ alleles were obtained with an average of 3.4 alleles per locus. For var. *hondurensis*, the average number of alleles per population was $n = 22$, ranging from 18 (H1) to 25 alleles (H6); the number of effective alleles was $A_e = 2.6$ and allelic richness was $A_r = 2.8$ (table 1). The diversity estimates for the other two varieties were lower than those obtained for var. *hondurensis*: $n = 19$ in var. *bahamensis* and $n = 20$ in var. *caribaea*, with values of $A_e = 2.37$ and $A_r = 2.7$ for both varieties (table 1). Two unique alleles were observed, one in population H4 of var. *hondurensis* (locus PtTX2146, 159pb) and another in population B2 of var. *bahamensis* (PtTX3025, 263pb). The average estimates of expected (H_e) and observed (H_o) heterozygosity were very similar for var. *hondurensis* (H_e = 0.575, H_o = 0.502) and var. *bahamensis* H_e = 0.576, H_o = 0.512), whereas the average estimates for var. *caribaea* were lower (H_e = 0.564, H_o = 0.476), although not statistically different from those of the other two varieties ($P > 0.05$). Three populations of var. *hondurensis* presented the highest genetic diversity values: population H6 from Guatemala (H_e = 0.634) and populations H10 and H11 from Nicaragua (H_e = 0.621 and H_e = 0.616, respectively). The lowest values were found in two populations of var. *hondurensis*, H1 (H_e = 0.488) and H9 (H_e = 0.553) sampled in Mexico and Honduras, respectively, and in population C2 of var. *caribaea* (H_e = 0.555) from Cuba (table 1). All the H_e values were higher than those of H_o (except in population B1 from the Bahamas), with a significant heterozygosity deficit of two or more loci per population in 9 of the 11 populations of var. *hondurensis*, the two populations (C1 and C2) of var. *caribaea* and population B2 of var. *bahamensis*. Most of the populations therefore contained fewer heterozygotes than expected under mutation-drift equilibrium.

The average inbreeding index for the varieties was positive and differed significantly from random mating expectations (F_{IS} = 0.131, P = 0.05, 95% CI: 0.043–0.203). *Pinus caribaea* var. c*aribaea* displayed the highest value, F_{IS} = 0.184 (P = 0.001, 95% CI: 0.025–0.349), followed by var. *hondurensis*, F_{IS} = 0.124 (P = 0.001, 95% CI: 0.034–0.225) and var. *bahamensis*, F_{IS} = 0.097 (P = 0.09, 95% CI: -0.130–0.223) (table 1). Most populations of var. *hondurensis* had significant levels of inbreeding, with the exception of two populations distributed in Belize (H3 and H4), and one in Honduras (H9). Population C2 of Mil Cumbres, var. *caribaea*, had the highest inbreeding index of all populations, F_{IS} = 0.319 (P = 0.001), together with B2 of var. *bahamensis* F_{IS} = 0.280 (P = 0.001). The other population of var. *bahamensis* (B1) presented a negative inbreeding index (F_{IS} = -0.063), but this was not significant. In general, most populations showed different degrees of inbreeding, indicating population isolation throughout their evolution.

Demographic reduction of population size considering the intermediate TPM mutation model showed eight out of 15 populations with signals of a bottleneck (deviation from mutation-drift equilibrium). The extreme models IAM and SMM, showed 13 populations and one population respectively (table 2). Based only on the TPM mutation model, populations H10 (Waspam) from Nicaragua, H7 (Mezapa) from Honduras, H2 (Carmelitas) and H3 (Rock) from Be-

Table 2 – **Bottleneck tests estimated for the three *Pinus caribaea* varieties distributed in the Caribbean Basin**.
The *T2*, bottleneck statistic (Cornuet & Luikart 1996) and *P*-values of the Wilcoxon signed rank test (one tail for heterozygosity excess) under the IAM, TPM and SMM mutation models, are shown for each population.

Population	*T2*, statistic					
	IAM	*P*-value	TPM	*P*-value	SMM	*P*-value
H1-Caobas	1.278	0.039	0.685	0.218	0.015	0.421
H2-Carmelitas	2.158	0.015	1.653	0.015	1.046	0.023
H3-Rock	2.027	0.007	1.410	0.015	0.778	0.053
H4-Mountain Pine	2.729	0.023	1.328	0.078	0.202	0.421
H5-DeepRiver	1.83	0.218	1.143	0.218	0.225	0.218
H6-Dolores	2.016	0.007	1.237	0.039	0.361	0.343
H7-Mezapa	2.218	0.015	1.607	0.039	0.915	0.078
H8-Trinidad	1.802	0.023	1.198	0.054	0.406	0.343
H9-Leimus	1.52	0.023	0.766	0.078	-0.341	0.656
H10-Waspam	2.345	0.007	1.715	0.007	0.896	0.078
H11-Moss	1.598	0.023	0.782	0.078	-0.315	0.578
C2-Mil Cumbres	1.967	0.007	1.415	0.023	0.776	0.218
C1-Viñales	2.018	0.015	1.422	0.039	0.705	0.078
B1-Andros	1.098	0.078	0.392	0.421	0.322	0.500
B2-New Providence	2.495	0.007	1.893	0.015	1.314	0.053

lize and H6 (Dolores) from Guatemala of var. *hondurensis*, showed heterozygosity excess indicating recent bottlenecks ($P < 0.05$). In island populations, the two Cuban populations were significantly bottlenecked, along with one population of Bahamas islands (B2-New Providence). These results clearly suggested that most populations of the three varieties had reduced population sizes in the recent past.

Genetic relationships between populations and varieties

The best clustering solution of populations obtained with BAPS comprised K=7 groups (*log*ML = -3612; $P = 1.000$) (fig. 1, electronic appendix 1). The first five groups comprised populations of var. *hondurensis*, the variety was thus spatially and genetically sub-structured: the $K1$ group consisted of two populations (H3 and H5 from Belize), the $K2$ group had four populations (H2 from Belize, H8, H9 from Honduras and H10 from Nicaragua), the $K3$ and $K4$ groups were represented by only one population each (H1 from Mexico; H7 from Nicaragua) and the $K5$ group included three populations (H4 from Belize, H6 from Guatemala, and H11 from Nicaragua). The $K6$ group included the two populations of var. *caribaea* (C1 and C2 from Cuba), and the $K7$ group the two populations of var. *bahamensis* (B1 and B2 from the Bahamas). In this way, we observed a tendency of the populations to cluster together in accordance with varieties and geographical distribution. The AMOVA analysis indicated a significant differentiation between varieties as estimated with R_{ST} (R_{ST} or $F_{CT} = 0.088$, $P < 0.001$; 95% CI: 0.024–0.103), which was of similar magnitude to differentiation between populations ($R_{ST} = 0.082$, $P < 0.001$; 95% CI: 0.064–0.132). The highest variance was found within individuals ($F_{IT} = 0.178$; 95% CI: 0.110–0.307), followed by variance between individuals within populations of each variety ($F_{IS} = 0.058$; 95% CI: 0.045–0.123) and between populations within varieties ($F_{SC} = 0.043$; 95% CI: 0.025–0.129). All of the values were significant ($P < 0.001$). Similarly, the AMOVA between the seven groups obtained with the BAPS analysis reflected hierarchical population structure ($R_{ST} = 0.077$; $P < 0.002$; 95% CI: 0.042–0.109) (table 3, electronic appendix 2A–C).

The neighbour-joining tree based on Nei's standardized genetic distance (D_a) showed two large groups supported by 100% of the bootstraps (fig. 2). The first comprised seven populations of var. *hondurensis*, indicating that the populations distributed in Belize (H3 and H5) were the most basal, while the most derived populations were H9 from Honduras and H10 from Nicaragua (fig. 2). The second group included the remaining populations of this variety and those of the varieties *caribaea* and *bahamensis*. The most basal distant

Table 3 – R_{ST} Fixation indices obtained with AMOVA analysis on three levels of grouping of *Pinus caribaea* varieties.
In bold type the highest value of F_{ST} obtained. Statistical significance was obtained with 1000 non-parametric permutations (Excoffier & Lischer 2010).

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	R_{ST} Fixation index	*P*-value
Among populations	14	8882.512	14.4117	8.2907	0.0829	< 0.001
Among groups of varieties	2	3963.351	16.1625	8.8417	**0.0884**	< 0.001
Among *K*7 groups	6	6477.252	13.5866	7.7547	0.0775	< 0.002

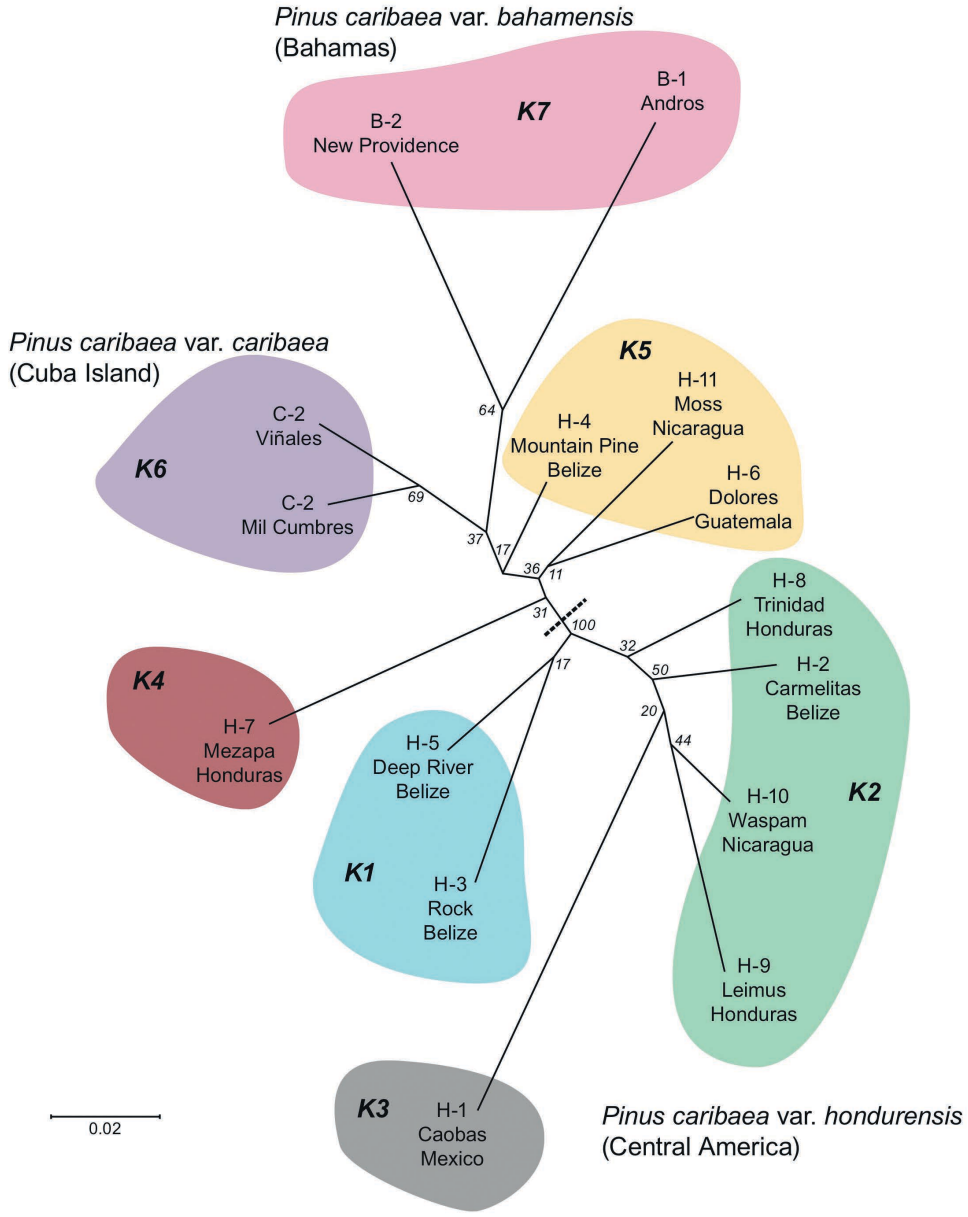

Figure 2 – Distance tree using neighbour-joining method, based on standardized genetic distances *Da* (Nei et al. 1983), between fifteen populations representing all three varieties of *Pinus caribaea*. Bootstrap support values of 1000 replicates are indicated at the base of the branches (in italics). Coloured clusters represent the *K* seven population groups obtained with Bayesian analysis (BAPS). Dashed line indicates the origin of the two large lineages (clusters) obtained.

Table 4 – **Migration rates (*M*) and historical effective size (N_e) estimates in *K* = 7 groups of *Pinus caribaea* varieties obtained with BAPS analysis.**
In bold type the highest values of *M* obtained between pairs of *K* groups. Readings of the values of *M* from left (row) to right (column). Markov chains were obtained with 10 000 burn-in steps and 500 000 iterations (Beerli 2008). CI-95%, Confidence intervals are given in brackets.

M (migration rate)	*K*1	*K*2	*K*3	*K*4	*K*5	*K*6	*K*7	θ ($4N_e\mu$)	N_e Historical
P. caribaea var. *hondurensis* *K*1 (H3 and H5)	0.0	2.10	1.67	3.03	2.27	**9.18**	**4.00**	1.2924 [0.5333–3.0666]	323.17 [132.22–766.67]
*K*2 (H2, H8, H9, H10)	**4.38**	0.0	3.64	1.53	0.77	3.45	0.84	1.7731 [1.0000–3.5333]	443.27 [250.00–883.33]
*K*3 (H1)	1.50	1.40	0.0	0.45	1.02	3.53	0.47	0.7989 [0.0000–2.4666]	200.72 [0.00–616.66]
*K*4 (H7)	1.74	**4.16**	1.40	0.0	1.01	1.71	1.71	1.2172 [0.4000–2.9333]	304.31 [100.00–733.33]
*K*5 (H4, H6, H11)	**8.32**	**4.57**	**7.96**	**10.90**	0.0	**20.16**	**5.56**	2.1463 [1.1333–4.3333]	537.00 [288.33–1083.32]
P. caribaea var. *caribaea* *K*6 (C1 and C2)	1.26	1.30	**7.54**	2.02	3.43	0.0	2.37	0.6456 [0.0000–2.2666]	161.41 [0.00–566.66]
P. caribaea var. *bahamensis* *K*7 (B1 and B2)	2.63	1.52	0.970	**4.11**	1.10	**4.85**	0.0	0.8328 [0.0000–2.5333]	208.19 [0.00–633.33]

populations were those of the var. *hondurensis*, distributed in Honduras (H7), Guatemala (H6) and Nicaragua (H11). The island populations of var. *caribaea* and var. *bahamensis* represented derived populations within two independent subgroups (fig. 2).

Effective size, migration rates and isolation by distance

The average estimate of historical effective population size (N_e) for the seven BAPS groups of *P. caribaea* was 362 individuals. The highest N_e was for group *K*5 of var. *hondurensis* (N_e = 537), comprising two populations from Belize (H4, H6) and one from Nicaragua (H11), followed by group *K*2 of the same variety (N_e = 443; H2, H8, H9 and H10) (table 4). Group *K*6 formed by two populations of var. *caribaea* (C1 and C2 of Cuba), had the lowest value (N_e = 161 individuals), followed by var. *hondurensis* (N_e = 201) and var. *bahamensis* (N_e = 208) from *K*3 (H1, Mexico) and *K*6 (B1 and B2, Bahamas) groups, respectively. Migration rate estimates between pairs of BAPS groups were comprised between *M* = 0.47 and *M* = 20.16. Inferred migration was predominantly from the continent to the islands, departing from groups *K*5 and *K*1 (*M* = 20.16 and 9.18) of var. *hondurensis*, and within the continent between the populations of group *K*5 towards groups *K*1 (*M* = 8.32), *K*2 (*M* = 4.57), *K*3 (*M* = 7.96), and *K*4 (10.90) of the same variety (see table 4 and fig. 3). Low *M* values were observed between the island varieties despite their relative geographic proximity (*M* = 4.85). The highest *M* value of var. *caribaea* was found towards population H1 of the *K*3 group (Mexico) of var. *hondurensis* (*M* = 7.54). Thus, we observed that the highest migration rate estimates were associated with dispersion from the distribution range of var. *hondurensis*. This indicates that *P. caribaea* possibly originated on the continent.

The IBD analysis among all of the populations showed a significant relationship between genetic and geographic distances (r^2 = 0.263; *P* = 0.005) (fig. 4), whereas no association was detected between differentiation among the seven BAPS groups and geographic distance (r^2 = 0.215; *P* = 0.139), or among the populations of var. *hondurensis* and geographic distance (r^2 = -0.065, *P* = 0.611).

DISCUSSION

Genetic variation and demographic history

We showed that populations of the three varieties of *Pinus caribaea* studied here had intermediate levels of genetic diversity when compared to other pine species studied with nSSRs (Williams et al. 2000, Shepherd et al. 2002, Dvorak et al. 2009, Sánchez et al. 2014, Zinck & Rajora 2016, Budde et al. 2017). The mean number of alleles (*A* = 3.4), allelic richness (A_r = 2.7–3.2) and effective number of alleles (A_e = 2.1–2.6) were similar to earlier findings in *P. caribaea* var. *bahamensis* (A_r = 3.2, Sánchez et al. 2014) and lower than values obtained for other pine species, such as *P. resinosa* Aiton (*A* = 9.0, Boys et al. 2005), *P. pinaster* Aiton (A_r = 8.3–10.2, A_e = 3.1–4.2, Mariette et al. 2001; *A* = 9.8, A_r = 9.2, De-Lucas et al. 2009), *P. taeda* (*A* = 4.9, Williams et al. 2000; *A* = 5.4, Al Rabah'ah & Williams 2004) and *P. oocarpa* (A_r =11.86, Dvorak et al. 2009). Diversity levels in *P. caribaea* were similar to those found in *P. halepensis* (A_r = 2.6–3.8, Budde et al. 2017). These comparisons should be taken with caution due to variation in sample sizes, marker polymorphism and completeness of geographic range sampled. Also, H_e is more robust than A_r because H_e is less affected by marker polymorphism. In this sense, the mean H_e across all populations (H_e = 0.571) was within the reported range for pine species (0.341 to 0.800, Karhu 2001, De-Lucas et al. 2009, Zinck & Rajora 2016, Budde et al. 2017), and was very similar to that obtained in *P. taeda* (H_e = 0.520, Al Rabah'ah & Williams 2004) and *P. patula* (H_e = 0.586, Dvorak et al. 2009). Populations of the var. *hondurensis* distributed in Guatemala (H6) and Nicaragua (H10 and H11) had the highest values

of H_e, whereas H3 of Mexico, and B1 of Bahamas Islands displayed the lowest values. However, for some of the populations, H_o was lower than expected under mutation–drift equilibrium, suggesting non-random mating.

In fact, the inbreeding levels of nine populations were significant (table 1). Inbreeding has been more drastic for the population Mil Cumbres (C2) of var. *caribaea*, which showed the highest value (F_{IS} = 0.319). This population is restricted to a small area in Cajálbana in western Cuba (70 km^2), growing only on ferrous serpentine soils (Marrero et al. 1998). Similarly, the Dolores (H6) population of var. *hondurensis*, located in Guatemala, presented high inbreeding (F_{IS} = 0.157). This small population is restricted to savanna forest and represents the most westerly distribution of this variety. Also, the population New Providence (B2) of var. *bahamensis*, the smallest forest area of all the Bahamas Islands, presented a high F_{IS} value (0.280). From this population, a study with five nSSRs also showed a significant F_{IS} (0.090) and a low H_e (0.487) (Sánchez et al. 2014). This population has lost around 64% of its initial extension due to deforestation and urbanization during the last century (Sánchez 2012), which could have led to a loss of genetic diversity with a consequent increase in inbreeding. In contrast, other populations with small size did not show significant inbreeding. For example, Caobas (H1) a small stand in Mexico surrounded by tropical semi-perennial forest and located in the northernmost distribution of var. *hondurensis*, displayed the lowest value of H_e (0.488) and did not significantly deviate from Hardy-Weinberg equilibrium (F_{IS} =0.057). A previous work using six nSSRs (four of them were used in this study) showed a similar value of H_e (0.471), but the inbreeding coefficient was higher and significant (F_{IS} = 0.097, $P <$ 0.05; Delgado et al. 2011). These differences could be due to sample size differences or marker choice; in this work the estimations were obtained based on 17 individuals whereas the work of Delgado et al. (2011) used 60 individuals. It has been suggested that this population might be a remnant one (Dvorak et al. 2005, Delgado et al. 2011). Another particular example is the northeast population of Andros Island (B1) of the var. *bahamensis*, where the F_{IS} was not significant (-0.063). This result is similar to that obtained by Sánchez et al. (2014) for the same population (F_{IS} = 0.019) and another seven populations studied in the Bahamas Islands (0.063 to -0.063). Also, population B1 did not show evidence of a recent population size decline (see table 2). Therefore, these results suggest that these populations from the Bahamas are in demographic equilibrium, where long distance gene flow through pollen dispersal, soil seed bank and wind-dispersed

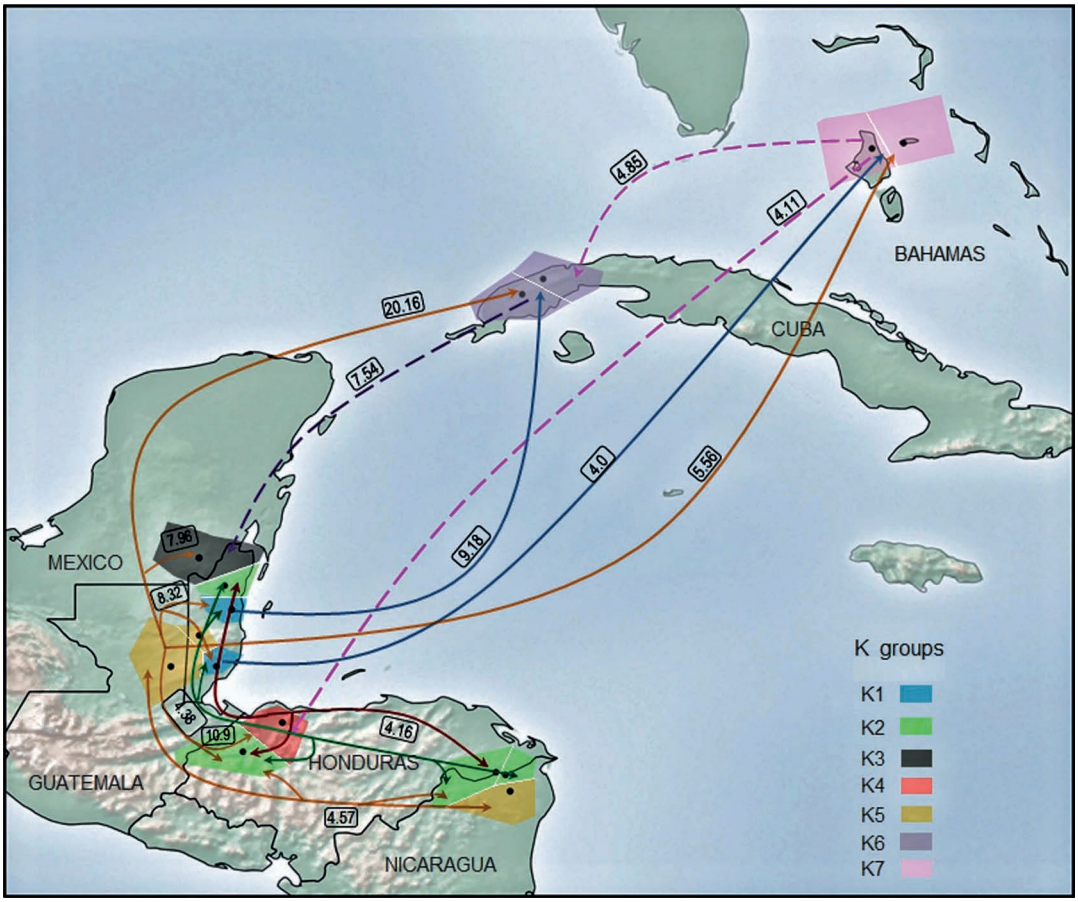

Figure 3 – Diagrammatic map with the high dispersal routes of the fifteen populations for *Pinus caribaea* varieties, obtained with Migrate analysis among the *K* seven groups defined by Bayesian analysis (BAPS). The figure shows only the migration rates (*M*) greater than four units, indicated inside of rectangles. Coloured lines represent the dispersal routes; continuous lines are the migration routes from continent-continent and continent-island groups. Dashed lines represent the dispersal routes among islands and islands-continent.

seeds of scattered mature individuals could have contributed to the maintenance of genetic variation (Sánchez et al. 2014).

In contrast to the previous example, the results of bottleneck tests obtained with two of the three mutation models (IAM and TPM) supports the hypothesis that most populations of the *P. caribaea* varieties showed signals of recent population bottlenecks, where allele number is reduced faster than heterozygosity (Cornuet & Luikart 1996). Several populations with isolated distribution and/or small N_e, presented the highest deviation of genetic diversity from expectations under demographic equilibrium (T_2), such as H10 from Nicaragua, in the southernmost part of the distribution of var. *hondurensis*; H6, a fragmented population from Guatemala, the two populations from Cuba, and B2, the New Providence population from the Bahamas (table 2). Our results indicate that most *P. caribaea* populations have experienced a historic bottleneck in their effective population size due to fragmentation and geographical isolation. Recent processes of colonization could be also plausible, at least from the island populations with significant signals of bottlenecks. Events of colonization have also been demonstrated with the use of cpSSR, from the same island pine varieties (Jardón-Barbolla et al. 2011).

The historical effective population size according to the groups obtained with the BAPS analysis was higher for *P. caribaea* var. *hondurensis* in which a total of five genetic clusters were observed (each with N_e of between 201 and 537 individuals) than for var. *bahamensis* and var. *caribaea* which each harbored a single genetic cluster (N_e = 208 and N_e = 161, respectively). Also, the N_e estimates within clusters (gene pools) of var. *hondurensis* were heterogeneous, reflecting the degree of population fragmentation or isolation. In pines, higher N_e estimates are associated with high values of genetic diversity and larger census population sizes (Ledig 1998, Rajora et al. 2000, Delgado et al. 2002, Ma. et al. 2006, Naydenov et al. 2014). For example, in *P. densata* Masters, an ancestral hybrid species with a large distribution in the Tibetan Plateau, the estimated N_e with seven loci was 73 200 (Ma et al. 2006). In contrast, in *P. pinaster*, a species with fragmented populations distributed in the Mediterranean region, a small N_e of 86.8 was obtained using eight nSSR (range of 42.5–359.1), suggesting signals of a demographic decline due to a recent bottleneck (Naydenov et al. 2014). In the same sense, for *P. resinosa*, with a larger and fragmented distribution in the northern USA and southern Canada, a small N_e of 142 was estimated (range of 62–222), using four nSSR and the same N_e estimator as this study (Beerli 2008), probably caused by an extreme genetic bottleneck (Boys et al. 2005). These values of N_e are more similar to those obtained for the island varieties of *P. caribaea* with a restricted distribution and some marginal land populations of var. *hondurensis*, most of them showing signals of recent bottlenecks.

Genetic relationships between populations and varieties

The analyses performed to assess the genetic structure of the populations of *P. caribaea* identified the varieties as a significant level of genetic differentiation. The average value of R_{ST} among the varieties was 0.088, indicating that 8.8% of the genetic variation is distributed among varieties. This value

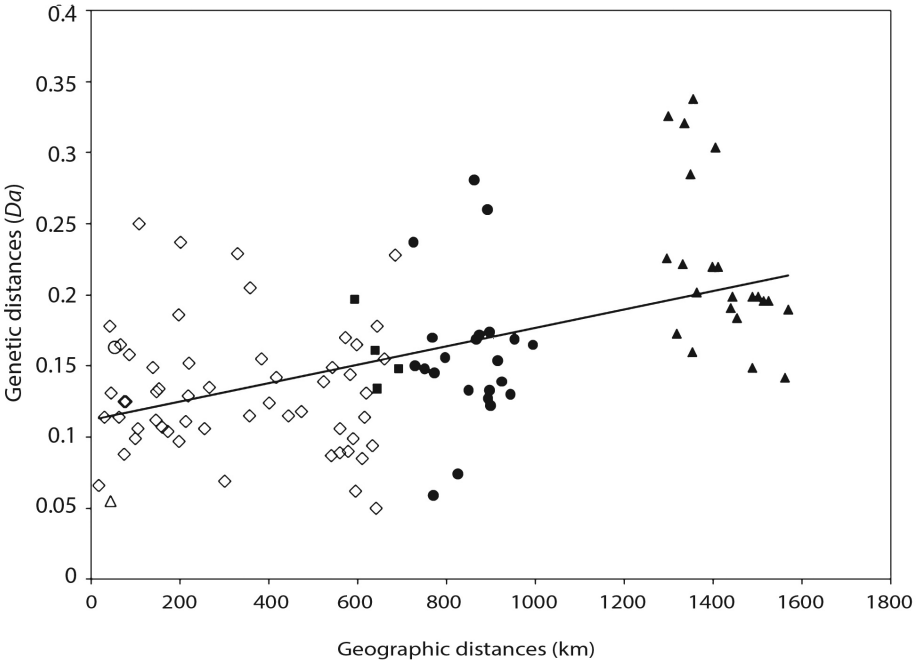

Figure 4 – Pattern of isolation by distance (IBD) among *Pinus caribaea* population pairs distributed in the Caribbean Basin. The correlation value was low but significant (r^2 = 0.263; P = 0.005), where 26% of the observed differences on the genetic distance can be attributed to geographical distance between populations. P-value was obtained with 10 000 permutation using Mantel test. Black symbols indicate the association among population pairs of varieties *hondurensis*/*caribaea* (●), *hondurensis*/*bahamensis* (▲), and *caribaea*/*bahamensis* (■). Open symbols indicate the association within var. *hondurensis* (◊), var. *caribaea* (○) and var. *bahamensis* (Δ) populations.

was relatively low, but lay within the values obtained for other pine species with nSSR (*P. pinaster*, R_{ST} = 0.111, Mariette et al. 2001; *P. resinosa*, R_{ST} = 0.280, Boys et al. 2005; *P. radiata* D.Don, R_{ST} = 0.145, Karhu et al. 2006; *P. taeda*, R_{ST} = 0.041, Al-Rabab'ah 2003; *P. oocarpa*, R_{ST} = 0.130, *P. tecunumani*, R_{ST} = 0.075 and *P. patula* Schltdl. & Cham., R_{ST} = 0.083, Dvorak et al. 2009) and was in fact higher than those obtained for some studied populations of *P. caribaea* var. *hondurensis* (R_{ST} = 0.021, Furlan et al. 2007; R_{ST} = 0.033, Delgado et al. 2011). While it may not be convenient to compare different markers, studies based on isoenzyme variation indicated weak genetic structure among the populations of the varieties *caribaea* and *hondurensis* (F_{ST} = 0.020, Zheng & Ennos 1999, and F_{ST} = 0.023, Dvorak et al. 2005, respectively), and moderate genetic structure in var. *bahamensis* (F_{ST} = 0.078, Zheng & Ennos 1999). In the present study, the highest genetic differentiation was found among populations of var. *hondurensis* (F_{ST} = 0.085), followed by var. *bahamensis* (F_{ST} = 0.076), with an F_{ST} similar to the one found by Zheng & Ennos (1999), whereas weaker genetic differentiation was obtained among populations of var. *caribaea* (F_{ST} = 0.059) (electronic appendix 2D–F). More recently, a phylogeographic study of the subsection *Australes* obtained higher levels of genetic differentiation for the three varieties using plastid microsatellites (cpSSRs) (R_{ST} = 0.230; Jardón-Barbolla et al. 2011). Since cpSSRs are haploid markers, they are of course more susceptible to the effects of genetic drift, and the absence of recombination in cpDNA does not obscure the geographic structure associated to gene genealogies as may be the case for nSSRs. It is therefore expected to obtain higher values of R_{ST}, generating a more marked genetic differentiation (Rosenberg & Nordborg 2002, Petit et al. 2005, Avise 2009).

The results obtained with the Bayesian analysis of population structure (BAPS) and the tree topology were dependent on variety, distinguishing the populations of the two insular varieties (group *K*6 of the var. *caribaea* and *K*7 of the var. *bahamensis*) from the populations of var. *hondurensis*. This may suggest that the varieties represent three independent evolutionary lineages. The populations of var. *bahamensis* (B1 on New Providence and B2 on Andros Island) are separated by a few kilometres (53.16 km), and the two populations of var. *caribaea* are located in the northern Cuba, at a relatively short distance apart (44.27 km). Geographical distance therefore does not seem to have played a major role in the genetic structure within each variety and each of these two varieties conforms to a specific genetic cluster. In contrast, populations of var. *hondurensis* show genetic structure within their geographical range; populations H9 and H10, located in the southern extreme of the distribution in Central America (Honduras and Nicaragua), are the most derived, while the populations H6 of Guatemala, H11 of Nicaragua and H7 of Honduras, are closer to the other varieties. These results are comparable to those obtained by Jardón-Barbolla et al. (2011), using cpSSRs, in which a marked phylogeographic structure was obtained, since haplotypes were not shared among the three varieties and a more significant relationship between the haplotypes of the varieties *bahamensis* and *hondurensis* was found. The latter variety also had substructure between its populations that comprised two groups; group I was distributed in the north (various populations of var. *hondurensis* and the two populations of the var. *bahamensis*), and group II in southern Central America. In this study, the substructure obtained for var. *hondurensis* was greater, given that five genetic groups were present. This is most likely due to the larger population size (2N) of nSSR relative to cpSSR (N), providing information from both progenitors (pollen and ovules). Isolation of groups could therefore be related to poor movement of seeds and/or pollen between some of the populations of these taxa.

Isolation by distance and dispersal routes

The IBD analysis showed a moderate correlation between genetic and geographic distances among all populations of the three varieties, indicating that nearby populations are less genetically isolated from each other than populations from different regions. However, when the analysis was conducted on populations of var. *hondurensis* alone, IBD was not significant, which once again suggests that the genetic differences between the populations of this variety are due to ecological factors. In this sense, a regional metapopulation dynamic has been stated by Jardón-Barbolla et al. (2011), where some populations are of recent formation while others tend towards extinction (Slatkin 1977). Populations that do not adjust to the IBD model are located in different areas of the species distribution. For example, the population Moss (H11) in southern Nicaragua has high values of R_{ST} with its neighboring population Leimus (H9; R_{ST} = 0.136) of Honduras, and is very similar to other, more geographically distant, populations; in the genetic distance tree, it clusters with Mountain Pine (H4) and Dolores (H6) from the north, and with Mezapa (H7) distributed in central Honduras (fig. 2). This latter population shares the highest number of alleles with populations distributed in the northern and southern regions; in the tree topology it is located in the early diverging part of the second group and thus could represent one of the most ancestral populations. The population Caobas (H1), distributed in the northern region, presents high values of differentiation with the populations Rock (H3; R_{ST} = 0.157) and Deep River (H5; R_{ST} = 0.091) distributed at a close distance (Belize) and is clustered with two of the southern populations (H10; Waspam and H9; Leimus). Furthermore, these populations have low values of H_e (H1, 0.421; H5, 0.491), and high values of inbreeding (H5, 0.154; H11, 0.177). These results give partial support to the hypothesis of a metapopulation dynamic (Jardón-Barbolla et al. 2011), since some populations located on the periphery or coastal lowland of the distribution area of this variety contain the lowest values of genetic variability and the smallest N_e (e.g. H1, Caobas in Mexico). Other populations located in the center of the distribution area of the variety (most of which exceed 300 individuals) contain alleles that are representative of the gene pool of the species and have a large N_e (e.g. H4, Mountain Pine, Belize or H7, Mezapa, Honduras). Therefore, the possible evolutionary scenario of var. *hondurensis* could be associated with expansion and contraction events of its populations. In particular, populations included in the *K5* group (H11 and H4/H6) could support this hypothesis: these populations are geographically distant, though genetically similar, which could be explained by a metapopulation dynamic, in which some

populations retain ancestral allelic variants (probably in process of expansion) while others do not (due to population contraction or extinction). This scenario is likely considering that pine savannas are frequently subjected to forest fires (Dvorak et al. 2005, Jardón-Barbolla et al. 2011).

Currently, as explained in the introduction, there are two general hypotheses regarding the dispersal routes of pine species in the Caribbean Basin. The first postulates that the initial migration involved an ancestor from Southern Florida to the Caribbean Islands (Adams & Jackson 1997), while the second hypothesis proposes that this dispersion could have occurred from Central America to the islands of Cuba and the Antilles (Mirov 1967, Dvorak et al. 2000a, 2005, Jardón-Barbolla et al. 2011). The results obtained in this study concur with the second hypothesis; the tree topology shows a closer association between the populations of var. *hondurensis* distributed in the central and southern regions of Central America and the two island varieties. The populations of var. *caribaea* and var. *bahamensis*, are nested within continent populations, suggesting that the dispersion initiated from Central America (Honduras, Nicaragua and Guatemala) towards the islands. The Migrate analysis corroborates these results and indicates that the main dispersal routes depart from the *K*5 and *K*1 groups from Central America towards all the populations of var. *hondurensis* and to those of the two island varieties (*K*6 and *K*7) with *M* values of between 4.57 and 20.16. The latter value is the migration rate obtained from var. *hondurensis* (*K*5) toward var. *caribaea* (*K*6) (table 4 and fig. 3). Three additional migration routes were suggested: (1) from the populations of var. *bahamensis* (*K*7 of the Bahamas) to var. *caribaea* (*K*6; *M*= 4.85, in Cuba), (2) from var. *bahamensis* (*K*7) to var. *hondurensis* (*K*4; *M* = 4.11, Mezapa, Honduras) and (3) from var. *caribaea* (*K*6) to var. *hondurensis* (*K*3; *M* = 7.54, Caobas, Mexico). These results support the hypothesis of demographic processes of expansion and contraction of population of var. *hondurensis*, and recent colonization to the two islands (Jardón-Barbolla et al. 2011), as well as sporadic events of migration from the islands to Central America.

Coalescence times of the Caribbean pine genetic clusters

Fossil-based divergence times calculated by Krupkin et al. (1996), as well as a recent phylogenetic reconstruction using chloroplast sequences calibrated with the fossil record (Hernández-León et al. 2013), suggested that the *Australes* group separated from its ancestors (*Oocarpae*) approximately 10 to 12 Mya; whereas Willyard et al. (2007), with evidence from nuclear and chloroplast loci and calibration with the fossil record, suggested a wider time interval (5 to 18 Mya). These divergence times may not be entirely correct; however, they do serve as a point of reference indicating that the ancestral clade of *P. caribaea* separated before or during this time. Taking into consideration the average value of the parameter θ in each gene pool obtained in the Migrate analysis, where $\theta = 4N_e\mu$ for diploid DNA (Hartl & Clark 1997), and where, mutation rate was assumed as 10^{-3} (Boys et al. 2005), the expected coalescence time within each gene pool can be obtained from N_e data of the table 4 as $2N_e$. Estimates vary between 323 generations for *P. caribaea* var. *caribaea* (*K*6) to 1074 for one of the gene pools of *P. caribaea* var. *hondurensis* (*K*5). Considering time to first reproduction between 10 (Okoro 1984) and 15 years (Jardón-Barbolla et al. 2011), one could consider a generation time of c. 30 years. The mean coalescence time within clusters would thus be on the order of between 9700 and 32 000 years, bearing in mind wide confidence intervals associated to the highly stochastic coalescent process. Divergence between varieties should a priori precede within-cluster coalescence, which is congruent with a speciation time of *P. caribaea* dated to the later Pliocene or early Pleistocene based on chloroplast and nuclear sequences data (Hernández-León et al. 2013). Since our coalescence time estimates for genetic clusters were recent, we inferred that demographic processes detected within clusters probably affected the last tens or hundreds of generations. Evidence from pollen records in Guatemala (Petén) have shown that the forests in the region included pines, oaks and elms, along with certain rainforest elements that were dated to between 8000 and 7000 years before present (Leyden 1984, Dvorak et al. 2005). Therefore, the distribution and abundance of the var. *hondurensis* in this region certainly expanded and contracted along with climatic changes over the last 10 000 years (Dvorak et al. 2005), which is partially consistent with the chronological times obtained in this study and the bottleneck detected for some populations of this complex of *P. caribaea*.

CONCLUSIONS

Our results support the hypothesis of the recent origin of this taxon from an ancestor of Central America (Honduras); the inferred migrations were predominantly from the continent to islands with sporadic migration events from the islands to continent. Thus, we deduce that the greatest source of genetic diversity is Central America, the area of distribution of var. *hondurensis*. Moreover, similar values of genetic diversity and shared genetic variation between the three varieties indicate that their speciation is not yet concluded. Most populations of var. *hondurensis*, and one of the two populations of each island variety, showed significant levels of inbreeding, with the highest levels found in those populations with marginal and coastal lowland distribution and small N_e. The historical demography of this species could be associated with long distance colonization events, followed by expansion and contraction of their populations.

SUPPLEMENTARY DATA

Supplementary data are available in pdf at *Plant Ecology and Evolution*, Supplementary Data Site (http://www.ingentaconnect.com/content/botbel/plecevo/supp-data) and consist of the following: (1) summary of Bayesian BAPS results of three Caribbean pine varieties with the *K*-groups generated, and (2) AMOVA results for different groups of data from three Caribbean pine varieties.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the comments and suggestions of David Gernandt, Keith MacMillan, the editor Myriam Heuertz and two anonymous reviewers that greatly

improved the quality of the manuscript. We also thank the Leon University Herbarium of Nicaragua, Pinar del Río University of Cuba, Belize Wildlife Service, Honduras University Herbarium (TEFH) and Santo Domingo Botanical Garden, F. Chi, P. Simá, R. Balam, N. Soltero, D. Escudero and Gretel Geada, for their help and cooperation in the collection of samples. We are grateful to G. Castillo, J. Coello, and A. Quiróz for their assistance with molecular work, and to Silvia Hernández-Aguilar and Germán Carnevali, for providing information of CICY herbarium database and registration of some samples. This work was financed by the Doctoral scholarship CONACYT-203335 to V. Rebolledo, and the projects CONACYT-44373 to P. Delgado and CONACYT-46925 to D. Piñero.

REFERENCES

Adams D.C., Jackson F.J. (1997) A phylogenetic analysis of the southern pines (Pinus subsecc. Australes Loudon): biogeographical and ecological implications. Proceedings of the Biological Society of Washington 110: 681–692.

Al-Rabab'ah M.A. (2003) Evolutionary dynamic of Pinus taeda L. in the late Quaternary: an interdisciplinary approach. PhD thesis, Texas A & M University, College Station, Texas, USA.

Al-Rabab'ah M.A., Williams C.G. (2004) An ancient bottleneck in the Lost Pines of central Texas. Molecular Ecology 13: 1075–1084. https://doi.org/10.1111/j.1365-294X.2004.02142.x

Avise J.C. (2009) Phylogeography: retrospect and prospect. Journal of Biogeography 36: 3–15. https://doi.org/10.1111/j.1365-2699.2008.02032.x

Beerli P. (1998) Estimation of migration rates and population sizes in geographically structured populations. In: Carvalho G.R. (ed.) Advances in Molecular Ecology: 39–54. IOS Press, Amsterdam.

Beerli P. (2008) Migrate version 3.4.2, a maximum likelihood and Bayesian estimator of gene flow using the coalescent. Available at http://popgen.sc.fsu.edu/oldversions/3.x/3.4/ [accessed 16 Nov. 2017].

Beerli P., Palczewski M. (2010) Unified framework to evaluate panmixia and migration direction among multiple sampling locations. Genetics 185: 313–326. https://doi.org/10.1534/genetics.109.112532

Bohonak A.J. (2002) IBD (Isolation by Distance): a program for analysis of isolation by distance. Journal of Heredity 93: 153–154. https://doi.org/10.1093/jhered/93.2.153

Boys J., Cherry M., Dayanandan S. (2005) Microsatellite analysis reveals genetically distinct populations of red pine (Pinus resinosa, Pinaceae). American Journal of Botany 92: 833–841. https://doi.org/10.3732/ajb.92.5.833

Budde K.B., González-Martínez S.C., Navascués M., Burgarella C., Mosca E., Lorenzo Z, Zabal-Aguirre M., Vendramin G.G., Verdú M., Pausas J.G., Heuertz M. (2017) Increased fire frequency promotes stronger spatial genetic structure and natural selection at regional and local scales in Pinus halepensis Mill. Annals of Botany 119: 1061–1072. https://doi.org/10.1093/aob/mcw286

Corander J., Sirén J., Arjas E. (2008) Bayesian spatial modeling of genetic population structure. Computational Statistics 23: 111–129. https://doi.org/10.1007/s00180-007-0072-x

Cornuet J.-M., Luikart G. (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001–2014.

Dakin E.E., Avise J.C. (2004) Microsatellite null alleles in parentage analysis. Heredity 93: 504–509. https://doi.org/10.1038/sj.hdy.6800545

Delgado P., Cuenca A., Escalante A.E., Molina-Freaner F., Piñero D. (2002) Comparative genetic structure in pines: evolutionary and conservation consequences. Revista Chilena de Historia Natural 75: 27–37.

Delgado P., Piñero D., Rebolledo V., Jardón L., Chi F. (2011) Genetic variation and demographic contraction of the remnant populations of Mexican Caribbean pine (Pinus caribaea var. hondurensis: Pinaceae). Annals of Forest Science 68: 121–128. https://doi.org/10.1007/s13595-011-0013-2

De-Lucas A.I., González-Martínez S.C., Vendramin G.G., Hidalgo E., Heuertz M. (2009) Spatial genetic structure in continuous and fragmented populations of Pinus pinaster Aiton. Molecular Ecology 18: 4564–4576. https://doi.org/10.1111/j.1365-294X.2009.04372.x

Dvorak W.S., Jordon A.P., Hodge G.P., Romero J.L. (2000a) Assessing evolutionary relationships of pines in the Oocarpae and Australes subsections using RAPD markers. New Forests 20: 163–192. https://doi.org/10.1023/A:1006763120982

Dvorak W.S., Gutiérrez E.A., Hodge G.R., Romero J.L., Stock J., Rivas O. (2000b) Pinus caribaea var. hondurensis. In: CAMCORE Cooperative (eds) Conservation and testing of tropical and subtropical forest tree species: 12–33. North Carolina State University.

Dvorak W.S., Hamrick J.L., Gutiérrez E.A. (2005) The origin of Caribbean pine in the seasonal swamps of the Yucatán. International Journal of Plant Science 166: 985–994. https://doi.org/10.1086/449314

Dvorak W.S., Potter K.M., Hipkins V.D., Hodge G.R. (2009) Genetic diversity and gene exchange in Pinus oocarpa, a Mesoamerican pine with resistance to the pitch canker fungus (Fusarium circinatum). International Journal of Plant Science 170: 609–626. https://doi.org/10.1086/597780

Echt C.S., May-Marquardt P., Hseih M., Zahorchak R. (1996) Characterization of microsatellite markers in eastern white pine. Genome 39: 1102–1108. https://doi.org/10.1139/g96-138

Eckenwalder J.F. (2009) Conifers of the world: the complete reference. Portland (OR), Timber Press.

El Mousadik A., Petit R.J. (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [Argania spinosa (L.) Skeels] endemic to Morocco. Theoretical and Applied Genetics 92: 832–839. https://doi.org/10.1007/BF00221895

Elsik G.C., Minihan T.V., Hall E.S., Scarpa M.A., Williams G.C. (2000) Low-copy microsatellite markers for Pinus taeda L. Genome 43: 550–555. https://doi.org/10.1139/g00-002

ESRI (1992–2000) ArcView version 3.2. Redlands, CA.

Excoffier L., Lischer H.E. (2010) Arlequin suite version 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10: 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Farjon A. (1996) Biodiversity of Pinus (Pinaceae) in Mexico: speciation and palaeo-endemism. Botanical Journal of the Linnean Society 121: 365–384. https://doi.org/10.1111/j.1095-8339.1996.tb00762.x

Farjon A., Styles B.T. (1997) Pinus (Pinaceae). Flora Neotropical Monograph 75. New York, New York Botanical Garden.

Farjon A. (2005) Pines, drawings and descriptions of the genus Pinus. 2nd Ed. Leiden & Boston, Brill.

Farjon A. (2013) Pinus caribaea var. hondurensis. The IUCN Red List of Threatened Species 2013: e.T20364366A20405060. https://doi.org/10.2305/IUCN.UK.2013-1.RLTS.T20364366A20405060.en

Flores L.C., López U.J., Vargas H.J.J. (2005) Indicadores reproductivos en poblaciones naturales de Picea mexicana Martínez. Agrociencia 39:117–126.

Francis J. K. (1992) Pinus caribaea Morelet. Caribbean pine. SO-ITF-SM-53. USDA Forest Service, Southern Forest Experiment Station. New Orleans & Los Angeles, Institute of Tropical Forestry.

Furlan R. de A., Mori E.S., Tambarussi E.V., Bueno de Moraes C., de Jesús F.A., Zimback L. (2007) Estrutura genética de populacões de melhoramento de Pinus caribaea var. hondurensis por meio de marcadores microssatélites. Bragantia 66: 553–563. https://doi.org/10.1590/S0006-87052007000400004

Gernandt D.S., Geada López G., Ortiz García S., Liston A. (2005) Phylogeny and classification of Pinus. Taxon 54: 29–42. https://doi.org/10.2307/25065300

Gernandt D.S., Pérez-de la Rosa J.A. (2014) Biodiversidad de Pinophyta (coníferas) en México. Revista Mexicana de Biodiversidad Suppl. 85: S126–S133. https://doi.org/10.7550/rmb.32195

Guo W.S., Thompson A.E. (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48: 361–372. https://doi.org/10.2307/2532296

Hartl D.L., Clark A.G. (1997) Principles of population genetics. 3rd Ed. Sunderland (MA), Sinauer Associates.

Hedges S.B. (1996a) The origin of West Indian amphibians and reptiles. In: Powell R., Henderson R.W. (eds) Contributions to West Indian herpetology: a tribute to Albert Schwartz: 9–128. Ithaca, Society for the Study of Amphibians and Reptiles.

Hedges S.B. (1996b) Historical biogeography of West Indian vertebrates. Annual Review of Ecology and Systematics 27: 163–196. https://doi.org/10.1146/annurev.ecolsys.27.1.163

Hernández-León S., Gernandt D.S., Pérez de la Rosa J.A., Jardón-Barbolla L. (2013) Phylogenetic relationships and species delimitation in Pinus Section Trifoliae inferred from plastid DNA. Plos One 8: e70501. https://doi.org/10.1371/journal.pone.0070501

Iturralde-Vinent M.A. (2004–2005) La Paleogeografía del Caribe y sus implicaciones para la biogeografía histórica. Revista del Jardín Botánico Nacional 25–26: 49–78.

Iturralde-Vinent M.A. (2006) Meso-Cenozoic Caribbean paleogeography: implications for historical biogeography of the region. International Geology Review 48: 791–827. https://doi.org/10.2747/0020-6814.48.9.791

IUCN (2016) The IUCN Red List of Threatened Species, version 2016.3. Available from http://www.iucnredlist.org [accessed 11 May 2017].

Jardón-Barbolla L., Delgado-Valerio P., Geada-López G., Vázquez-Lobo A., Piñero D. (2011) Phylogeography of Pinus subsection Australes in the Caribbean Basin. Annals of Botany 107: 229–241. https://doi.org/10.1093/aob/mcq232

Karhu A. (2001) Evolution and applications of pine microsatellites. Oulu, Oulu University Press.

Karhu A., Vogl C., Moran G.F., Bell J.C., Savolainen O. (2006) Analysis of microsatellite variation in Pinus radiata reveals effects of genetic drift but no recent bottlenecks. Journal of Evolutionary Biology 19: 167–175. https://doi.org/10.1111/j.1420-9101.2005.00982.x

Krupkin A.B, Liston A., Strauss S.H. (1996) Phylogenetic analysis of the hard pines (Pinus subgenus Pinus, Pinaceae) from chloroplast DNA restriction site analysis. American Journal of Botany 83: 489–498.

Ledig F.T. (1998) Genetic variation in Pinus. In: Richardson (ed.) Ecology and biogeography of Pinus: 251–280. Cambridge, Cambridge University Press.

Leyden B.W. (1984) Guatemalan forest synthesis after Pleistocene aridity. Proceedings of National Academy of Sciences of the United States of America 81: 4856–4859.

Ma X.F., Szmidt A.E., Wang X.R. (2006) Genetic structure and evolutionary history of a diploid hybrid pine Pinus densata inferred from the nucleotide variation at seven gene loci. Molecular Biology and Evolution 23: 807–816. https://doi.org/10.1093/molbev/msj100

Mantel N. (1967) The detection of disease clustering and a generalized regression approach. Cancer Research 27: 209–220.

Mariette S., Chagné D., Lézier C., Pastuszka P., Raffin A., Plomion C., Kremer A. (2001) Genetic diversity within and among Pinus pinaster populations: comparison between AFLP and microsatellite markers. Heredity 86: 469–479.

Marrero A., Renda A., Calzadilla E. (1998) Comportamiento de Pinus caribaea var. caribaea Morelet en diferentes tipos de suelos. Revista Cuba Forestal 1: 39–40.

Matheson A.C., Bell J.C., Barnes R.D. (1989) Breeding systems and genetic structure in some Central America pine populations. Silvae Genetica 38: 107–113.

Millar C.I. (1993) Impact of the Eocene on the evolution of Pinus L. Annals of Missouri Botanical Garden 80: 471–498. https://doi.org/10.2307/2399795

Mirov N.T. (1967) The genus Pinus. New York, The Ronald Press Company.

Moonlight P.W., Richardson J.E., Tebbitt M.C. Thomas D.C., Hollands R. Peng C.-I., Hughes M. (2015) Continental-scale diversification patterns in a megadiverse genus: the biogeography of Neotropical Begonia. Journal of Biogeography 42: 1137–1149. https://doi.org/10.1111/jbi.12496

Naydenov K.D., Alexandrov A., Matevski V., Vasilevski K., Naydenov M.K., Gyuleva V., Carcaillet C., Wahid N., Kamary S. (2014) Range-wide genetic structure of maritime pine predates the last glacial maximum: evidence from nuclear DNA. Hereditas 151: 1–13. https://doi.org/10.1111/j.1601-5223.2013.00027.x

Nei M., Tajima F., Tateno Y. (1983) Accuracy of estimated phylogenetic trees from molecular data. II Gene frequency data. Journal of Molecular Evolution 19: 153–170. https://doi.org/10.1007/BF02300753

Nikles D.G. (1966) Comparative variability and relationships of Caribbean Pine (Pinus caribaea Mor.) and Slash Pine (Pinus elliottii Engelm.) PhD thesis, College of Natural Resources, North Carolina State University, Raleigh, North Carolina, USA.

Okoro O.O. (1984) Influence of flowering habit of Pinus caribaea var. hondurensis Barr. et Golf. on seed yield. Nigerian Journal of Forest 13: 36–42.

Perry J. (1991) The pines of Mexico and Central America. Portland, Oregon Timber Press.

Petit J.R., Duminil J., Fineschi S., Hampe A., Salvini D., Vendramin GG. (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Molecular Ecology 14: 689–701. https://doi.org/10.1111/j.1365-294X.2004.02410.x

Pindell J., Kennan L., Stanek K.P., Maresch W.V., Draper G. (2006) Foundations of Gulf of Mexico and Caribbean evolution: eight controversies resolved. Geologica Acta 4: 303–341.

Piry S., Luikart G., Cornuet J.-M. (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. Journal of Heredity 90:502–503. https://doi.org/10.1093/jhered/90.4.502

QGIS Development Team (2017) QGIS version 2.14.21. Available from https://www.qgis.org/en/site/ [accessed 2 Jul. 2017].

Rajora O.P., Rahman M.H., Buchert G.P., Dancik B.P. (2000) Microsatellite DNA analysis of genetic effects of harvesting in old-growth eastern white pine (Pinus strobus) in Ontario, Canada. Molecular Ecology 9: 339–348. https://doi.org/10.1046/j.1365-294x.2000.00886.x

Rosenberg N.A., Nordborg M. (2002) Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nature Reviews Genetics 3: 380–390. https://doi.org/10.1038/nrg795

Sánchez M. (2012) Conservation genetics and biogeography of the Caribbean pine (Pinus caribaea var. bahamensis) in the Bahamas archipelago. PhD thesis, Birkbeck University of London, UK.

Sánchez M., Ingrouille M.J., Cowan R.S., Hamilton M.A., Fay M.F. (2014) Spatial structure and genetic diversity of natural populations of the Caribbean pine, Pinus caribaea var. bahamensis (Pinaceae), in the Bahaman archipelago. Botanical Journal of the Linnean Society 174: 359–383. https://doi.org/10.1111/boj.12146

Schlötterer C. (2000) Evolutionary dynamics of microsatellite DNA. Chromosoma 109: 365–371. https://doi.org/10.1007/s004120000089

Schuchert C. (1935) Historic geology of the Antillean-Caribbean region. New York, John Wiley & Sons Inc.

Shepherd M., Cross M., Maguire T.L., Dieters M.J., Williams C.G., Henry R.J. (2002) Transpecific microsatellites for hard pines. Theoretical and Applied Genetics 104: 819–827. https://doi.org/10.1007/s00122-001-0794-z

Slatkin M. (1977) Gene flow and genetic drift in a species subject to frequent local extinctions. Theoretical Population Biology 12: 253–262. https://doi.org/10.1016/0040-5809(77)90045-4

Slatkin M. (1995) A measure of population subdivision based on microsatellite allele frequency. Genetics 139: 457–462.

Sokal R.R., Rohlf FJ. (1995) Biometry: the principles and practice of statistics in biological research. New York, W.H. Freeman & Co.

Szpiech Z.A., Jakobsson M., Rosenberg N.A. (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. Bioinformatics 24: 2498–2504. https://doi.org/10.1093/bioinformatics/btn478

Takezaki N., Nei M. (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics 144: 389–399.

Takezaki N., Nei M., Tamura K. (2010) POPTREE2: software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. Molecular Biology and Evolution 27: 747–752. https://doi.org/10.1093/molbev/msp312

Tsuda Y., Chen J., Stocks M., Lascoux M., Källman T., Sonstebo J. H., Parducci L., Semerikov V., Sperisen C., Politov D., Ronkainen T., Väliranta M., Vendramin G.G., Tollefsrud M.M., Lascoux M. (2016) The extent and meaning of hybridization and introgression between Siberian spruce (Picea obovata) and Norway spruce (P. abies): cryptic refugia as stepping stones to the west? Molecular Ecology 25: 277–2789. https://doi.org/10.1111/mec.13654

Van W.G. (2002) Pinus caribaea Morelet. In: CAB International (eds) Pines of silvicultural importance: 3–50. Wallingford, UK, CABI Publishing.

van Oosterhout C., Hutchinson W.F., Wills D.P.M., Shipley P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Vázquez-Lobo A. (1996) Evolución de hongos endófitos del género Pinus L.: implementación de técnicas moleculares y resultados preliminares. PhD Thesis, Facultad de Ciencias, UNAM, México.

Wang M.L., Barkley N.A., Jenkins T.M. (2009) Microsatellite markers in plants and insects. Part I: applications of biotechnology. Genes, Genomes and Genomics 3: 54–67.

Williams C.G., Elsik C.G., Barnes R.D. (2000) Microsatellite analysis of Pinus taeda L. in Zimbabwe. Heredity 84: 261–268.

Willyard A., Syring J., Gernandt D.S., Liston A., Cronn R. (2007) Fossil calibration of molecular divergence infers a moderate mutation rate and recent radiations for Pinus. Molecular Biology and Evolution 24: 90–101. https://doi.org/10.1093/molbev/msl131

Wright S. (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19: 395–420. https://doi.org/10.2307/2406450

Zheng Y.Q., Ennos R.A. (1999) Genetic variability and structure of natural and domesticated populations of Caribbean pine (Pinus caribaea Morelet). Theoretical and Applied Genetics 98: 765–771. https://doi.org/10.1007/s001220051133

Zinck JWE., Rajora OP. (2016) Post-glacial phylogeography and evolution of a wide-ranging highly-exploited keystone forest tree, eastern white pine (Pinus strobus) in North America: single refugium, multiple routes. BMC Evolutionary Biology 16: 1–17. https://doi.org/10.1186/s12862-016-0624-1

Manuscript received 8 Mar. 2017; accepted in revised version 12 Dec. 2017.

Communicating Editor: Myriam Heuertz.