


# Genetic diversity of *Enterolobium cyclocarpum* in Colombian seasonally dry tropical forest: implications for conservation and restoration

E. Thomas<sup>1</sup>  · C. Gil Tobón<sup>2</sup> · J. P. Gutiérrez<sup>3</sup> ·  
C. Alcázar Caicedo<sup>4</sup> · L. G. Moscoso Higueta<sup>5</sup> ·  
L. A. Becerra<sup>3</sup> · J. Loo<sup>6</sup> · M. A. González<sup>7</sup>

Received: 27 April 2016 / Revised: 6 November 2016 / Accepted: 2 December 2016  
© Springer Science+Business Media Dordrecht 2016

**Abstract** *Enterolobium cyclocarpum* is a characteristic legume tree species of seasonally dry tropical forests (SDTFs) of Mesoamerica and northern South America typically used in silvopastoral and agroforestry systems. Remaining populations of *E. cyclocarpum* in Colombia are severely fragmented owing to the highly degraded state of SDTF in the country, posing threats to both their in situ persistence and their usefulness as seed sources for future planting efforts. We genotyped *E. cyclocarpum* populations at nine sampling sites across a latitudinal gradient of SDTF in Colombia by means of eight nSSR markers to elucidate the species diversity distribution in the country. Our data suggest that a deep divide seems to have existed between Caribbean and Andean populations of *E. cyclocarpum* in Colombian SDTF that may date back to the last glacial maximum (~21,000 BP), or longer. However, we only found evidence of genetic differentiation between trees from the southern Cauca River valley and populations at more northern locations. All the latter populations showed signs of admixture which may be the result of human-influenced

---

Communicated by Daniel Sanchez Mata.

---

This article belongs to the Topical Collection: Forest and plantation biodiversity.

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s10531-016-1274-8) contains supplementary material, which is available to authorized users.

---

✉ E. Thomas  
evert.thomas@gmail.com

<sup>1</sup> Bioversity International, Lima, Peru

<sup>2</sup> Universidad de Antioquia, Medellín, Colombia

<sup>3</sup> International Center for Tropical Agriculture, Cali, Colombia

<sup>4</sup> Bioversity International, Cali, Colombia

<sup>5</sup> Forestpa SAS, Medellín, Colombia

<sup>6</sup> Bioversity International, Rome, Italy

<sup>7</sup> Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, Colombia

movement of germplasm, particularly after the introduction of cattle by European settlers. Most of the sampled sites showed heterozygosity scores close to Hardy–Weinberg expectations. Only the three southern-most populations displayed significantly positive values of inbreeding coefficient, potentially affecting their in situ maintenance and their use as seed sources. Based on our findings we identify priority areas for the in situ conservation of remaining *E. cyclocarpum* populations, and propose a strategy for sourcing of appropriate planting material for use in future tree planting efforts.

**Keywords** Paleodistribution · Agroforestry · Suitability modelling · *Enterolobium cyclocarpum* · Conservation · Fragmentation · Seed zones

## Introduction

*Enterolobium cyclocarpum* Griseb. is a characteristic legume tree species of seasonally dry tropical forests (SDTFs) of Mesoamerica and northern South America, one of the most endangered ecosystems worldwide (Miles et al. 2006). The species has supported human livelihoods throughout history. Its fruits, boiled unripe seeds and roasted ripe seeds are edible by humans, and its seeds can be ground into a protein-rich edible flour. Its wood is widely used for both timber and domestic applications, while its bark can be used for tanning leather and produces an exudate with similar characteristics as Arabic gum (Rocas 2002). Owing to the umbrella shape of its generally extensive canopy and the nutrient-rich fodder produced by its leaves and fruits, *E. cyclocarpum* is commonly used in silvopastoral and agroforestry systems (Lagemann and Heuvelink 1983; Escalante 1985; Mahecha 2002; Reinoso-Pérez 2014). This use is bound to continue in light of current restoration objectives established by countries associated with Initiative 20 × 20 which aims to initiate the restoration of 20 million hectares in Latin America and the Caribbean by 2020 (<http://www.wri.org/our-work/project/initiative-20x20>). Restoration in the context of Initiative 20 × 20 has a broad interpretation and includes the establishment of silvopastoral and agroforestry systems.

An important factor to ensure the long-term success of tree planting efforts refers to the quality of the planting material used. To bolster the viability and resilience of planted populations, reproductive material must be genetically diverse and matched to the current and expected future climate conditions of the planting sites (Breed et al. 2013; Thomas et al. 2014, 2015a). Continued availability of good quality germplasm of *E. cyclocarpum* also requires that adequate conservation measures are in place to ensure the in situ viability of identified seed sources.

The distribution and structure of genetic diversity in tree species with longstanding human use is often an outcome of the complex intertwining of their natural and human history (Thomas et al. 2012, 2015b; Galluzzi et al. 2015). An important aspect of a species' natural history relates to its response to past climate changes. It has been argued that SDTFs, the natural habitat of *E. cyclocarpum*, used to have more extensive and largely contiguous distributions during the late Pleistocene which may have reached their maxima during a cool and dry period ~18,000–12,000 BP, coinciding with the contraction of the humid forest. This is known as the 'Pleistocenic arc hypothesis' (Prado and Gibbs 1993; Pennington et al. 2000). In Colombia, the focus of this paper, outcomes of paleodistribution modelling are consistent with a more extensive distribution of SDTFs during the last

glacial maximum LGM (Werneck et al. 2012), which contracted to the isolated fragments known today as the consequence of fluctuations in climate during the Holocene (Marchant et al. 2002). Main SDTF fragments in Colombia are located in the Caribbean coastal area, specifically in the Cauca, Magdalena, Chicamocha and Patia river valleys (Pennington et al. 2000; Pizano et al. 2014). These fragments contain biological elements from Caribbean and Mesoamerican SDTFs in the north and from the Andes, Brazil and southern South America in the south (Pizano et al. 2014; Banda-R et al. 2016). If the formation of the current SDTF fragments in Colombia dates back to the early Holocene, it is likely that the tree populations that survived in these fragments have been isolated from each other long enough to have diverged genetically.

Today, most of the remaining *E. cyclocarpum* populations grow in areas bearing the scars of historical and ongoing anthropogenic disturbance. Significant forest clearance has taken place during the late Holocene period in Colombia (Marchant et al. 2001; Etter 2015) and intensified during the past century which was characterized by conversion of mature tropical forests to pastures for cattle ranching (Vina and Cavelier 1999). The regions most affected by degradation are the Caribbean and Andean regions which were home to the largest areas of SDTFs in Colombia (Etter and van Wyngaarden 2000; Etter et al. 2008; Etter 2015). Today, less than 8% of the original dry forest cover remains (García et al. 2014). In degraded landscapes one would expect to see footprints in the genetic profiles of remaining *E. cyclocarpum* populations as a result of multiple effects, such as domestic animal and human-mediated seed movement, habitat fragmentation and isolation of trees in pastures leading to a reduction in effective population sizes, among other impacts (Gonzales et al. 2010). Fragmentation is known to negatively affect the reproduction, gene flow and genetic diversity of tree populations (Lowe et al. 2005; Aguilar et al. 2006, 2008; de Abreu et al. 2015). Fragmentation may also shift mating patterns towards selfing, both of which, in combination with reduced effective population sizes, may increase the levels of inbreeding, with potentially detrimental effects on the vigor and viability of next generation populations (Rocha and Aguilar 2001; Aguilar et al. 2008).

Here we study the genetic diversity of *E. cyclocarpum* populations along a latitudinal gradient in Colombian SDTF. We assess whether paleodistribution modelling of *E. cyclocarpum* is concordant with the Pleistocenic Arc Hypothesis (Prado and Gibbs 1993; Pennington et al. 2000) and if any evidence of genetic differentiation can be detected between the sampled populations that can be related to paleoclimate-mediated isolation of suitable areas. Furthermore, we investigate potential signals of human disturbance and fragmentation on the genetic profiles of sampled populations. Based on our findings we identify priority areas for the in situ conservation of *E. cyclocarpum* populations and propose a pragmatic approach to guide the selection of appropriate planting material for use in tree planting efforts.

## Methods

### Field sampling

Between July 2014 and June 2015, we collected leaf material from 65 reproductive individuals of *E. cyclocarpum* at nine sampling sites across a latitudinal gradient (3–11° latitude) of SDTF in Colombia (2–16 trees per site; Table 1). Field sampling was guided by suitability maps based on occurrence data obtained from botanical collection records.

**Table 1** Genetic parameters of *Enterolobium cyclocarpum*, with and without bootstrap correction across all sampling sites

	N	Without resampling						With resampling (3 trees)					
		Na	LCA	I	Ho	He	F	Na	LCA	I	Ho	He	F
<b>ZAT</b>													
Mean	4	3.13	0.63	0.97	0.59	0.57	-0.05	2.69	0.50	0.86	0.59	0.52	-0.14
SD		0.83	0.74	0.27	0.29	0.11	0.47	0.84	0.64	0.32	0.33	0.15	0.54
<b>SENA</b>													
Mean	4	3.50	0.63	1.08	0.52	0.59	0.09	2.84	0.50	0.90	0.52	0.53	0.02
SD		1.20	0.74	0.48	0.35	0.25	0.47	1.05	0.64	0.44	0.38	0.24	0.55
<b>GUA</b>													
Mean	10	5.00	0.88	1.36	0.50	0.69	0.24	2.63	0.29	0.84	0.51	0.52	0.00
SD		2.45	0.99	0.52	0.24	0.15	0.35	1.01	0.51	0.38	0.35	0.18	0.58
<b>COL</b>													
Mean	3	3.50	0.75	1.07	0.56	0.59	0.09	3.50	0.75	1.07	0.56	0.59	0.09
SD		1.41	0.71	0.43	0.34	0.16	0.45	1.41	0.71	0.43	0.34	0.16	0.45
<b>ITU</b>													
Mean	10	4.38	0.25	1.21	0.40	0.63	0.33	2.64	0.07	0.83	0.40	0.50	0.17
SD		1.92	0.46	0.44	0.25	0.16	0.43	0.96	0.20	0.38	0.34	0.19	0.61
<b>COT-SFE</b>													
Mean	13	4.38	0.25	1.14	0.40	0.60	0.32*	2.40	0.08	0.73	0.40	0.46	0.10
SD		2.46	0.61	0.50	0.21	0.16	0.39	0.88	0.21	0.37	0.34	0.20	0.62
<b>PIN</b>													
Mean	16	4.38	0.38	1.05	0.35	0.55	0.35*	2.29	0.15	0.66	0.35	0.41	0.12
SD		2.45	0.52	0.54	0.25	0.21	0.36	1.00	0.28	0.43	0.35	0.24	0.57
<b>PAI</b>													
Mean	3	2.75	0.50	0.87	0.29	0.53	0.43*	2.75	0.50	0.87	0.29	0.53	0.43*
SD		0.89	0.76	0.33	0.21	0.15	0.44	0.89	0.76	0.33	0.21	0.15	0.44
<b>CIAT</b>													
Mean	2	2.88	0.5	0.94	0.63	0.55	-0.14	-	-	-	-	-	-
SD		1.23	0.53	0.49	0.44	0.26	0.54						

*N* sample size, *Na* Average allelic richness per locus, *LCA* average locally common allele richness per locus, *I* average Shannon index per locus, *Ho* average observed heterozygosity per locus, *He* average expected heterozygosity per locus, *F* average inbreeding coefficient per locus

*ZAT* Buffer zone of Tayrona National Park, *SENA* experimental station SENA, *GUA* El Guamo; *COL* Coloso, *ITU* Ituango, *COT-SFE* Cotové-Santa Fé, *PIN* Pintada, *PAI* La Paila, *CIAT* International center for tropical agriculture

\* Significantly different from 0 at  $p < 0.05$

Valle de Cauca department appears to be the south-western limit of the species' distribution in Colombia; our field sampling revealed that it does not grow naturally in the Patía valley, the southernmost area of SDTF in the country. Although the species is known to occur in the upper Magdalena and Chicamocha River valleys, we were unable to find any trees that were not clearly cultivated during field sampling missions to the departments of

Tolima, Huila and Santander. All biological material was collected by the Instituto Alexander von Humboldt following Decreto 302 of 2003.

## DNA extraction and PCR amplification

Leaf material was preserved in zip-lock bags with silica gel prior to processing in the laboratory. We extracted total genomic DNA from *E. cyclocarpum* leaves using the CTAB procedure (Doyle and Doyle 1990), with modifications based on Colpaert et al. (2005), Alzate-Marin et al. (2009) and Novaes et al. (2009). The polymerase chain reactions (PCRs) were carried out with 9 SSR primers (see Table S1 of the Electronic Supplementary Material 1 Table S1; Peters et al. 2008). The procedure for fluorescent dye labelling of PCR fragments in one reaction (Schuelke 2000) was performed with three primers: a sequence-specific forward primer with M13 (–19) tail at its 5' end, a sequence-specific reverse primer, and universal fluorescent-labeled M13 (–19) (–CAC-GACGTTGTAAAACGAC). 6-Fam, Vic and Ned fluorescent dyes were used for the chemical labelling of primers.

Polymerase chain reactions was performed in a total volume of 12  $\mu$ L containing 0.5 U of Platinum® Taq (Invitrogen®, USA), 1X PCR Buffer (200 mM Tris–HCl (pH 8.4), 500 mM KCl) (Invitrogen®, USA), 0.2 mM of each dNTP (Promega Corp., USA), 8.0 mg/mL BSA, 3.0 mM MgCl<sub>2</sub> (Invitrogen®, USA), 0.05 pmol/ $\mu$ L labelled M13 primer, 0.01 pmol/ $\mu$ L M13 tagged-Forward primer, 0.05 pmol/ $\mu$ L reverse primer and 40 ng of DNA. PCR was conducted in a Mastercycler® pro (Eppendorf, Germany) for 1 cycle at 95 °C during 2 min, followed by 15 cycles (30 s at 94 °C, 30 s at 65 °C and 30 s at 72 °C), followed by 35 cycles (15 s at 94 °C, 15 s at 50 °C and 45 s at 72 °C). The PCR products were separated in 2% agarose gel, to verify if the loci were polymorphic (Table S1). Primer Ency-21 was discarded because it had inconsistent results. Allele numbers of the retained loci varied between 3 (Ency8) and 16 (Ency24).

Polymerase chain reactions products were run on an ABI PRISM 3730 DNA Analyzer sequencer and sized with GeneScan-500LIZ (Applied Biosystems) size standard. Allele sizes were determined using the software GeneMapper version 4.0 (Applied Biosystems).

## Suitability modelling

We characterized the spatial distribution of favorable habitat for *E. cyclocarpum* in Colombia under different climatic conditions by means of suitability mapping based on ensembles of modelling algorithms, implemented in R package *BiodiversityR* (Kindt and Coe 2005). Habitat suitability during the LGM (~21,000 BP) and Mid Holocene (~6000 BP) was modeled to determine whether past climate conditions may have had an impact on the current distribution of genetic diversity. Habitat suitability under present and future climate conditions was modeled to assess the expected impact of climate change on the in situ conservation of *E. cyclocarpum* populations.

Presence data collected during field sampling were complemented with Colombian records extracted from numerous sources ([www.gbif.org](http://www.gbif.org); the national herbaria MEDEM, HUA y MEDEL; [www.dryflor.info](http://www.dryflor.info); [www.orinoquiabiodiversa.org](http://www.orinoquiabiodiversa.org); [www.sibcolombia.net](http://www.sibcolombia.net)). Only records from SDTF as defined by the combination of Etter et al. (2008) and García et al. (2014) were included in our dataset. As a result, 128 presence points were used for suitability modelling. Background points (an overall maximum of 10,000 and maximum one per grid cell) were randomly selected from the area enclosed by a convex hull polygon

constructed around all presence points and extended with a buffer corresponding to 10% of the polygon's largest axis.

We applied two different strategies for suitability modelling under past and future climate conditions. Model calibrations for projections to LGM and mid-Holocene climate conditions were carried out at 2.5 and 0.5 arc minute resolution, respectively, using only WorldClim climate layers (Hijmans et al. 2005) as explanatory variables. Model calibrations intended for projections to future climate scenarios (period 2040–2069; referred to as 2050 s) were carried out at 30 arc seconds resolution, using aside from climate layers also altitude, slope, aspect, terrain roughness, direction of water flow and seven major edaphic variables, obtained from ISRIC-World Soil Information (Hengl et al. 2014): organic carbon (ORCDRC), pH in H<sub>2</sub>O (PHIHOX), sand % (SNDPPT), silt % (SLTPPT), clay % (CLYPPT), Cation Exchange Capacity (CEC), Bulk density (BLD), Coarse fragments >2 mm (CRFVOL). For the edaphic variables we calculated a weighted mean across 0–5, 5–15, 15–30, 30–60, and 60–100 cm soil depth values in order to derive a single data value for 0–100 cm. Collinear explanatory variables were removed based on iterative calculations of variance inflation factors (VIF), retaining only variables with VIFs smaller than 5. The resulting sets of explanatory variables, as well as presence and background points used for model calibrations are given in Table S2 of the Electronic Supplementary Material 1.

Modelling algorithms considered in the ensembles were maximum entropy (MAX-ENT), boosted regression trees (BRT, including a stepwise implementation), random forests (RF), generalized linear models (GLM; including stepwise selection of explanatory variables), generalized additive models (GAM; including stepwise selection of explanatory variables), multivariate adaptive regression splines (MARS), regression trees (RT), artificial neural networks (ANN), flexible discriminant analysis (FDA), support vector machines (SVM), and the BIOCLIM algorithm. As spatial autocorrelation among species presence points is known to bias model evaluations based on cross-validation, we evaluated the ability of all individual modelling algorithms to cope with spatial autocorrelation by calculating calibrated Area Under Curve (cAUC) values and comparing these with a geographical null model (Hijmans 2012). We compared the cAUCs of each of the individual distribution models with the cAUCs of the geographical null model resulting from twenty iterations, by means of Mann–Whitney tests. Only models that gave cAUC values that were significantly higher than the null model were retained for the construction of different model ensembles. In a next step, we calculated the cAUC values for all possible ensemble combinations of the retained models. Each ensemble combination was constructed as the weighted average of its individual composing models, using their respective average cAUC values as weights. The ensemble that yielded the highest cAUC value was considered to generate the most appropriate scenario for projecting to past and future climate conditions, respectively.

To assess habitat suitability under mid-Holocene and LGM climate conditions we carried out projections to one (ECHAM3; Deutsches Klimarechenzentrum (DKRZ) 1992) and two (MIROC and CCSM; Braconnot et al. 2007) climate models, respectively. These same climate models were used by Werneck et al. (2011). For characterizing future climate conditions, we used 31 downscaled climate models for the period 2040–2069 based on the Representative Concentration Pathway (RCP) 4.5 scenario of greenhouse gas emissions, prepared for the Fifth Assessment IPCC report (CMIP5) (Ramírez Villegas and Jarvis 2010). We limited model projections to areas where suitability scores were higher than the maximum training sensitivity plus specificity threshold obtained from model calibration under current climate conditions. To obtain summarizing maps for the two LGM climate

models we averaged the threshold-limited suitability maps constructed for both individual climate scenarios. Two scenarios were considered for future suitability maps (period 2040–2069). Optimistic and pessimistic scenario maps were limited to areas which were identified as suitable by at least one and fifteen of all 31 threshold-limited climate projections, respectively. Model statistics are given in Table S2 of the Electronic Supplementary Material 1.

### Statistical analyses and diversity mapping

To visualize geographic patterns in nSSR diversity, we constructed circular neighborhoods of 10 arc minutes diameter ( $\sim 18$  km at the equator) at a spatial resolution of 30 arc seconds, following Thomas et al. (2012). These neighborhoods can be considered as proxies for each of the sampled populations. In overlap zones between neighborhoods genetic diversity parameters were calculated as if trees from both sampled populations were part of one effective breeding population. Grid-based calculations of genetic parameters included allelic richness, the Shannon information index, expected and observed heterozygosity, the inbreeding coefficient and the number of locally common alleles (LCA) per locus. LCA are alleles that occur only in a limited area of a species' distribution (here  $<25\%$  of the sampled populations) but reach high frequencies (here  $>5\%$ ) in those areas. High LCA richness can be indicative for the level of genetic isolation of populations (Frankel et al. 1995) and can hence be useful to identify locations of putative refugia (Thomas et al. 2012). To correct for sample bias, genetic parameters were calculated as the average values obtained for 1000 bootstrapped subsamples of the minimum sample size of 3 trees per grid cell. As only two trees were sampled at CIAT, this site was excluded from bootstrap-corrected genetic parameter calculations.

To visualize country-scale genetic gradients, we applied spatial principal component analysis (sPCA) (Jombart et al. 2008) implemented in *adegenet* package version 1.4–2 (Jombart 2008) for R. This method yields scores that summarize both the genetic variability and spatial structure among individuals. It uses both a matrix with allele frequencies of genotypes and a spatial weighting matrix containing measurements of spatial proximity among entities, based on a connection network. Here we used a distance-based neighborhood which is more appropriate for aggregated distributions (Jombart et al. 2008) as is the case for *E. cyclocarpum*. Minimum and maximum distances used were 0.1 and 2 arc degrees. More specifically, sPCA uses Moran's I to measure spatial structure of allele frequencies. Moran's I values are highly positive when allele frequencies observed at neighboring sites tend to be similar (contributing to global structures in data), whereas they are strongly negative when allele frequencies at neighboring sites are dissimilar (contributing to local structures). A sPCA generates two sets of axes: one set with positive eigenvalues and the other with negative eigenvalues. Positive eigenvalues correspond to global structures, while negative values are indicative of local patterns. Applied to the present dataset, a predominant local structure was detected, which means that neighboring trees tended to show stronger genetic differences than random pairs of trees. The negative eigenvalues of the last two axes were clearly much lower than all other eigenvalues and therefore we only interpreted the first local structure associated with the last two axes. This decision was confirmed by a Monte-Carlo test on the global and local structures in the dataset (simulated p values 0.01 and 0.09, respectively). We visualized the structure in the *E. cyclocarpum* dataset on a raster map with 30 s grid cells, by assigning to each cell the average value of the projections on the last sPCA axis (with most negative eigenvalue) of all individuals enclosed by a circular neighborhood of 10 arc minutes diameter constructed around its center.

We submitted our data to Bayesian cluster analysis using STRUCTURE (Pritchard et al. 2000) and adegenet's discriminant analysis of principal components (DAPC) (Jombart 2008). In STRUCTURE we tested three different modelling scenarios: an admixture ancestry model with and without consideration of sampling localities on the assumption that allele frequencies are correlated, and an admixture model with consideration of sampling localities on the assumption that allele frequencies are independent. The number of groups (K) tested varied between 1 and 9, using burnin periods of one million steps and 10 million additional replications. For each value of K, we carried out 10 independent repetitions. We used the method of Evanno et al. (2005) for detection of the most probable number of genetically homogeneous clusters (K), through calculation of  $\Delta K$  as implemented in the STRUCTURE HARVESTER software (Dent and VonHoldt 2011). DAPC was unable to detect any meaningful spatial structuring of the trees we sampled and the results obtained from STRUCTURE suggested a mixed origin of trees at most sampling sites. We therefore additionally used correspondence analysis implemented in *adegenet* (Jombart 2008) to visualize genetic grouping at population level. All maps were edited in ArcMap 10.2.

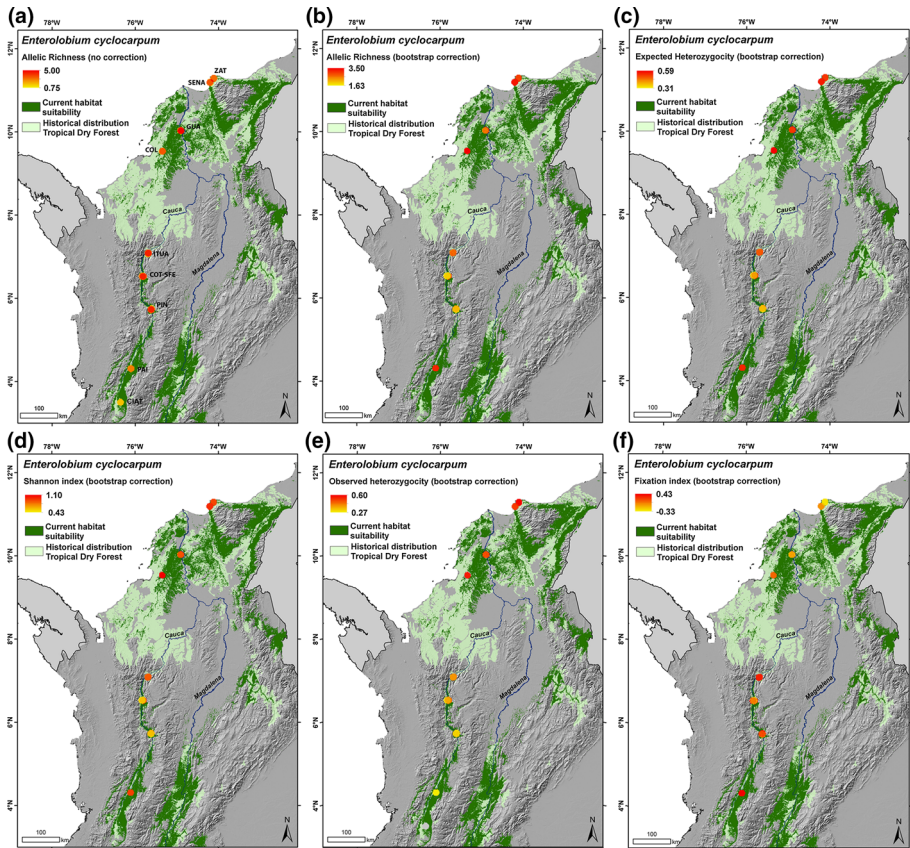
Complementary genetic analyses such as isolation by distance,  $F_{ST}$  (Nei 1973) and AMOVA (Excoffier et al. 1992) were carried out in R packages *adegenet* and *poppr* version 2.0.2 (Kamvar et al. 2014).

## Results

Average values and standard deviations of the main genetic parameters of all sampled *E. cyclocarpum* populations are given in Table 1. A comparison of the different parameters calculated with and without bootstrap correction shows the influence of sample bias on parameter scores, justifying the use of the bootstrap correction. Overall, different genetic diversity measures (allelic richness, Shannon diversity, expected heterozygosity) suggest that the most diverse populations are located in the Caribbean region (ZAT, SENA, COL, GUA) and the southern part of the Cauca River valley (PAI). Elevated values were also observed in Ituango (ITUA). The inbreeding coefficient varied from  $-0.33$  to  $0.43$ , but only values found for La Paila (PAI;  $F_{IS} = 0.43$ ), La Pintada (PIN) and Cotové-Santa Fé (COT-SFE) were significantly higher than zero ( $p < 0.05$ ; Table 1). Figure 1 shows the spatial distribution of the genetic diversity parameters against a background of the species' current habitat suitability and the historical distribution of SDTF in Colombia.

The modeled distributions of suitable habitat during past climates (Fig. 2) suggest that *E. cyclocarpum* populations from SDTF in Colombia may have had a much wider range during the LGM than at present and that current agglomerated distribution patterns may have formed in the early Holocene, as is evident from the strong overlap between suitable areas under current and mid-Holocene climate conditions (Figs. 1, 2c, d, respectively). Figure 2 also suggests that there may have been a deep division between populations of *E. cyclocarpum* from the Caribbean region and the Cauca River valley south of  $7.3^\circ$  latitude that dates back to the LGM, or longer. This division is mirrored by patterns in the distribution of LCA richness and scores on the last sPCA axis (Fig. 2). In the southern and central Cauca River valleys there was a trend of decreasing richness of locally common alleles from south to north, while in the Caribbean all values were at the higher end of allelic richness (Fig. 2a, c). Similarly, spatial representation of projections of sampled trees on the last sPCA axis (most negative eigenvalue) revealed a genetic gradient corresponding to the latitudinal gradient from the Caribbean region to the southern Cauca River valley, with the genetic cline being located in the middle Cauca River valley (Fig. 2b, d).

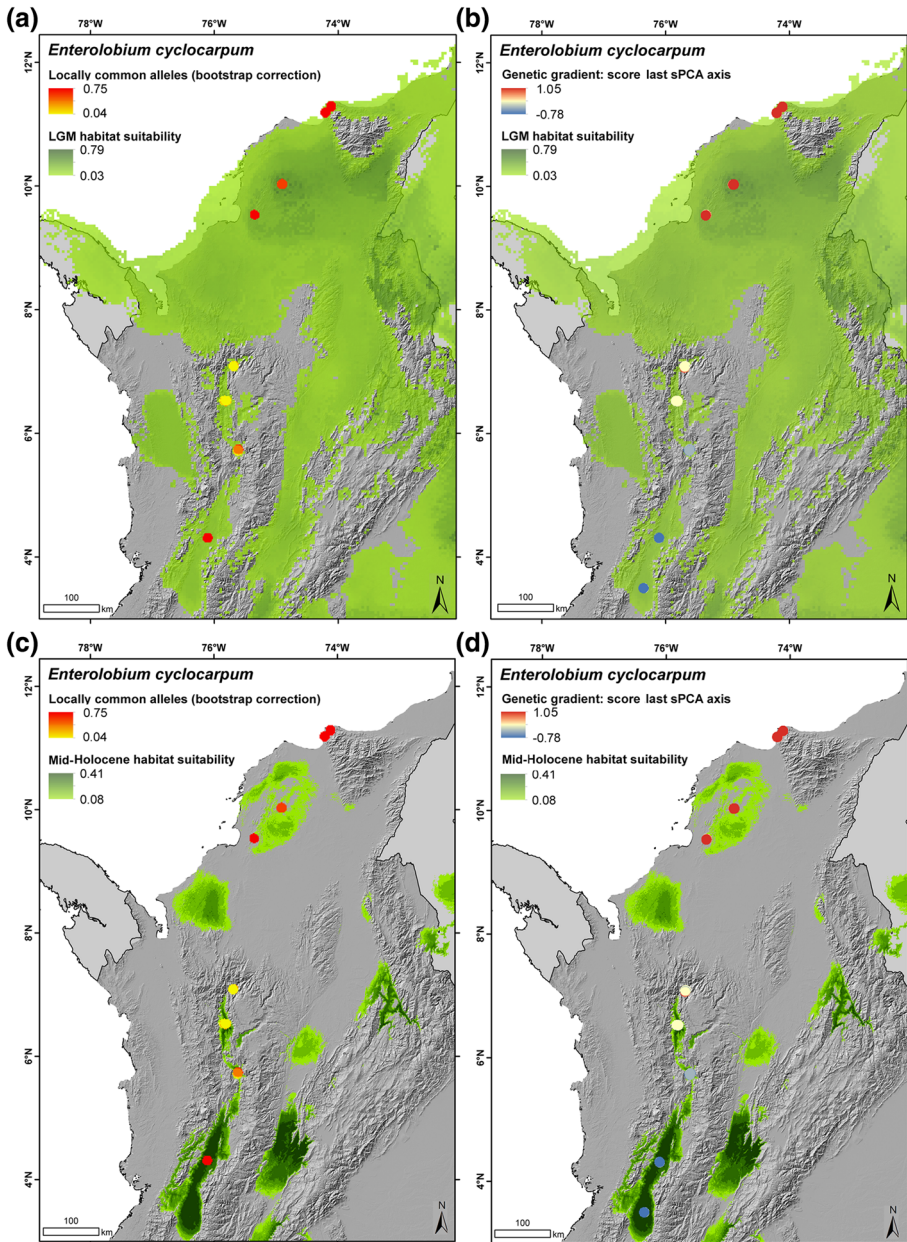




**Fig. 1** Spatial variation of different genetic parameters of *E. cyclocarpum* against a background of the species' current modeled habitat suitability, and the historical distribution of SDTF in Colombia: **a** allelic richness without sample bias correction; **b** allelic richness with bootstrap correction; **c** expected heterozygosity with bootstrap correction; **d** Shannon index with bootstrap correction; **e** observed heterozygosity with bootstrap correction; **f** inbreeding coefficient with bootstrap correction

While AMOVA indicated that only 6.6% of genetic variation in data resided between populations, compared to 93.4% between individuals within populations, pairwise  $F_{ST}$  values between sampling areas were in several cases high enough ( $>0.1$ ) to indicate some degree of genetic differentiation between populations (see Table S3 of the Electronic Supplementary Material 1). More specifically,  $F_{ST}$  values larger than 0.1 were consistently found between the trees sampled at la Paila (PAI) and all other sampled populations. Interestingly, relatively high  $F_{ST}$  values were also found between the trees sampled in the buffer zone of Tayrona national park (ZAT) and the nearby populations of Coloso (COL) and SENA, but not the populations sampled at larger distances.

Results of  $\Delta K$  computation showed support for  $K = 4, 3$  and  $2$  for the different modelling scenarios tested in STRUCTURE, the most geographically coherent being the admixture model with consideration of sampling locations and the assumption that allele frequencies are correlated. This scenario yielded support for both  $K = 2$  and  $3$  (Figure S1 of the Electronic Supplementary Material 2), of which  $K = 3$  was most informative, showing a clear separation between the sampling sites in the southern Cauca River valley

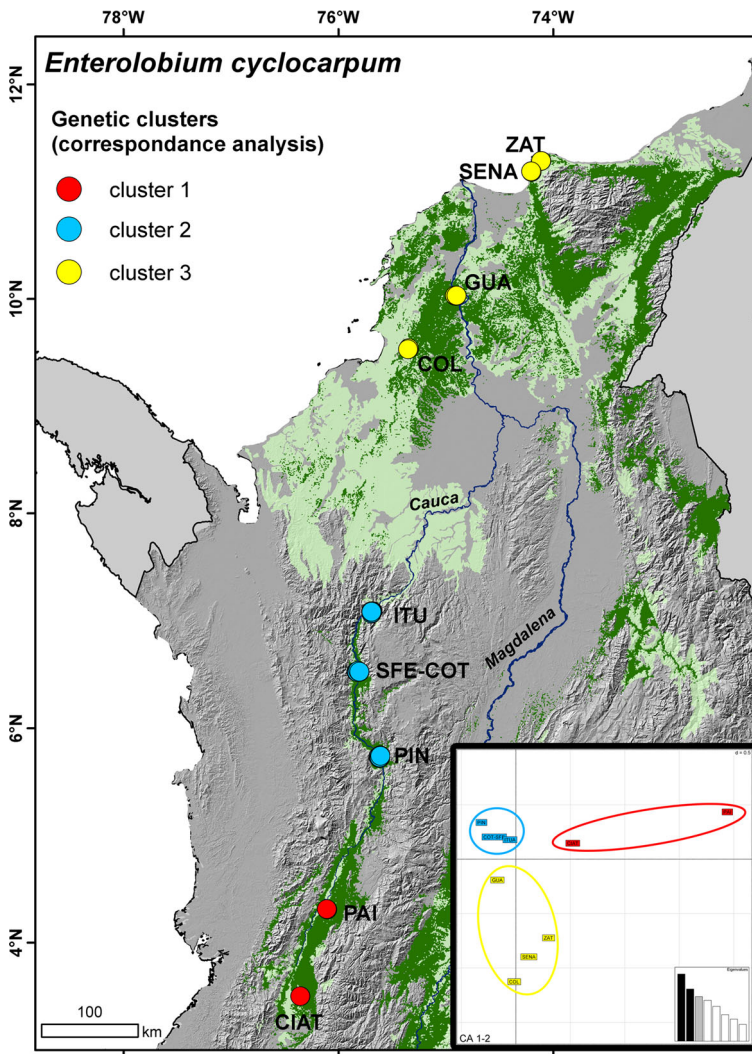


**Fig. 2** Modeled distribution of suitable habitat of *E. cyclocarpum* during the Last Glacial Maximum (LGM; ~21,000 BP) and the mid-Holocene (~6000 BP), in combination with richness of locally common alleles (**a** and **c**, respectively) and average scores of sampled trees as projected on the last sPCA axis (**b** and **d**, respectively)

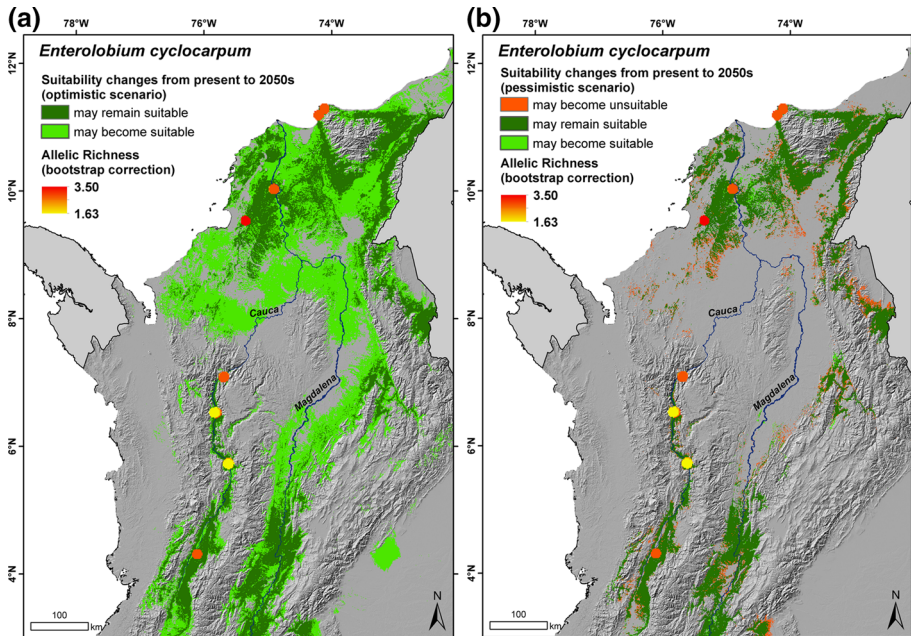
and all the other ones being composed of individuals assigned to two different clusters, corroborating the findings of pairwise  $F_{ST}$  comparisons. One of the two trees sampled at CIAT seems to have originated from germplasm from more northern sources.

Correspondence analysis also supported the genetic division between populations from the Caribbean region and the Cauca River valley and to a lesser extent between the southern and middle Cauca River valley (Fig. 3). This genetic structure, although weak, was corroborated by a Mantel test comparison of geographical and Nei's genetic distances between all possible pairs of *E. cyclocarpum* trees, which showed evidence of an isolation by distance pattern ( $p = 0.002$ ).

Future climate projections suggest that the expected impact of climate change on *E. cyclocarpum* habitat suitability is likely to be small. Even under a pessimistic scenario, all high diversity sites are likely to remain suitable in the near future, thus offering opportunities for in situ conservation of populations (Fig. 4).



**Fig. 3** Spatial representation of genetic groups of sampled populations of *E. cyclocarpum* based on correspondence analysis. The first and second axes explain 22 and 17% of the variation in data, respectively



**Fig. 4** Future (2050s) habitat suitability of *E. cyclocarpum* showing all areas identified as suitable by at least 1 (a) and 15 (b) of 31 future climate model projections, respectively, in combination with the distribution of allelic richness

## Discussion

Here we assessed the genetic diversity distribution of *E. cyclocarpum* populations along a latitudinal gradient of Colombian SDTF and how it may have been shaped by past climatic changes and more recent human influences. Our habitat suitability models for *E. cyclocarpum* during the LGM and mid-Holocene are consistent with the Pleistocenic Arc Hypothesis (Prado and Gibbs 1993; Pennington et al. 2000). While the current agglomerated distribution pattern of *E. cyclocarpum* in Colombia may have formed in the early Holocene (Fig. 2c, d) the isolation between populations from the Caribbean region and those located in the middle and southern Cauca River valleys may have started much earlier, during the LGM, or possibly even longer. This isolation hypothesis is partly supported by the results of our genetic analyses. First, the highest values of LCA richness—indicative of prolonged isolation—were found in all Caribbean populations and the southern Cauca River valley population of La Paila, respectively (Fig. 2c). Second, the results of the Bayesian cluster analysis (Figure S1b), the correspondence analysis (Fig. 3) and the  $F_{ST}$  comparisons (Table S3), all consistently separated the southern Cauca River valley population of La Paila (PAI) from all northern populations. And third, the fact that the genetic cline identified by our sPCA analysis is located in the middle Cauca River valley (Fig. 2d) underpins the possibility of long-term isolation of Caribbean and southern Cauca River valley populations of *E. cyclocarpum*.

The fact that LCA richness decreases from south to north in the southern to central Cauca River valley might suggest that expansion of *E. cyclocarpum* may have taken place in that direction at some point in the post-glacial history of the species. This is supported

by the low genetic diversity scores obtained for the otherwise best-sampled populations of Pintada (PIN) and Cotové-Santa Fé (COT-SFE) from the middle Cauca River valley (Fig. 1b–d). However, the genetic distinctness of these populations from the trees sampled at La Paila, might allude to the possibility of an alternative scenario whereby low-diversity *E. cyclocarpum* populations may have been isolated in the middle Cauca river valley in prehistoric times [e.g. around the population of La Pintada (PIN) where LCA richness is highest]. The LGM suitability map supports this scenario (Fig. 2a, b). Such isolation might have giving rise to one of the two clusters currently found in the middle Cauca river valley and the Caribbean region (Figure S1b), where the other cluster may have formed. After the European colonization, human-influenced movement of planting stock pertaining to these two clusters may have intensified, resulting in the admixed nature of most Caribbean and middle Cauca River valley populations (see below). Accordingly, elevated diversity observed in the Ituango (ITU) population at the northern end of the middle Cauca River valley might be the result of mixing between populations that migrated south from a Caribbean stock with populations that expanded northward from an origin in the southern or middle Cauca River valley (Petit et al. 2003).

Our genetic data are insufficient to support or refute the Pleistocenic Arc Hypothesis as our findings suggest that paleoclimate-induced isolation and genetic differentiation of populations may date back to the LGM (or longer), rather than to the Holocene. Testing this hypothesis will require a broader coverage and higher density of samples, including from important areas such as the Chicamocha and upper Magdalena river valleys, ideally combined with coalescence analysis based on different markers, including maternally inherited ones (Duminil et al. 2015).

In spite of the genetic differentiation between the southern Cauca River valley population of La Paila and all more northern populations, most of the genetic diversity resided within rather than among the populations we sampled, confirming previous work from Costa Rica (Gonzales et al. 2010). Low genetic divergence among populations from the middle Cauca River valley and the Caribbean region is consistent with *E. cyclocarpum*'s predominantly outcrossing mating system and its potential for long-distance pollen movement (Rocha and Lobo 1996) which is believed to be mediated by moths, beetles, and other small nocturnal insects (Janzen 1982), as well as diurnal insects such as bees (Rocha and Aguilar 2001). The extensive foraging ranges of some of these insects might result in significant movement of pollen among mass flowering trees, such as *E. cyclocarpum* (Rocha and Aguilar 2001).

Possibly more importantly, long-distance seed dispersal is likely to have contributed to the low genetic differentiation among populations. Janzen (1981, 1982) believed that *E. cyclocarpum*'s dispersal was originally mediated by now extinct Pleistocene horses and that at present most dispersal is carried out by domesticated horses and cows, and wild tapirs, peccaries and rodents. Due to *E. cyclocarpum*'s popularity as a shade tree in pastures and farm land since the initiation of the European colonization and possibly before, reproductive material has been distributed extensively through human intervention, either as seeds or seedlings, or as seeds in the gut of domestic animals. This could explain why sPCA revealed a predominant local structure, and why trees of mixed origin were identified in most sample sites by the Bayesian cluster analysis. Only in La Paila were all trees grouped in the same cluster, which suggests they had all originated from local seed sources. These trees had diameters at breast height between 235 and 305 cm, which could imply that at least the tallest one was more than 400 years old.<sup>1</sup> While recent introductions

<sup>1</sup> <http://www.monumentaltrees.com/en/trees/enterolobiumcyclocarpum/records/>.

of planting material of *E. cyclocarpum* from the Caribbean area to the southern Cauca River valley area are well known, it is quite unlikely that the trees we sampled in La Paila are the consequence of human interference. Cattle breeding in Colombia became significant only in the Seventeenth century (Etter 2015), implying that the cultivation of *E. cyclocarpum* for providing shade (and fodder) to cattle is unlikely to have been widespread at the probable time of establishment of these trees.

As expected for outcrossing, long-lived tree species, most of the sites we sampled yielded heterozygosity scores close to Hardy–Weinberg expectations, corroborating the findings of Gonzales et al. (2010). Only the three southernmost sampling sites of COT–SFE, PIN and PAI yielded significantly positive inbreeding coefficients. As inbreeding is one of the early signs of the effects of habitat fragmentation, while genetic diversity is lost more slowly over subsequent generations (Lowe et al. 2005), this could mean that the trees we sampled in these areas may already have been exposed to habitat fragmentation at the time of their establishment. Considering that the diameters of these trees varied between 35 and 340 cm, this may have been decades to centuries ago, in line with historical reconstructions of deforestation of Colombian SDTF (Marchant et al. 2001; Etter 2015). All other populations showed heterozygosity scores close to zero, possibly because their effective population sizes were larger. In a study from Guanacaste, Costa Rica, Gonzales et al. (2010) related the nonsignificant excess of heterozygous *E. cyclocarpum* adults in populations to the fact that reproductive trees were rarely separated by more than 1 km, a distance that is well within the documented range of pollen movement, or because the trees established prior to the manifestation of habitat fragmentation.

Elevated values of inbreeding coefficients are not only a concern for the in situ maintenance of the respective populations we sampled, but also for their use as seed sources for tree planting activities. When low effective population sizes persist over successive generations, a concomitant decrease of heterozygosity due to random drift and increase of inbreeding are expected. This in turn may result in the accumulation of deleterious recessive alleles, lowering the fecundity of individuals, increasing mortality of seeds and seedlings, and reducing the growth rate of individuals, eventually leading to population extinction (Young et al. 1996; Aguilar et al. 2008). Rocha and Aguilar (2001) found that while >80% of the progeny produced from single *E. cyclocarpum* trees in Costa Rica were sired by pollen from another parent, trees from continuous forest were almost six times more likely to set fruits and produce more seeds per fruit than trees in pastures. Furthermore, progeny from trees in continuous forest were more vigorous than those in pastures (Rocha and Aguilar 2001). Additionally, in pastures seed, seedling, and juvenile mortality has been reported to be high due to predation by small rodents, high susceptibility of seedlings to fire, desiccation, trampling, grazing, and competition with grasses. In the absence of human management intervention, the recruitment bottlenecks resulting from such processes are likely to fuel a self-reinforcing process eventually resulting in extinction of local populations (Janzen 1983; Rocha and Lobo 1996).

The recent commitments of Latin American countries to restore vast areas of degraded lands on their territories in the context of Initiative 20 × 20, provide opportunities to reverse the trend of forest degradation and fragmentation and improve the prospects for conservation of the remaining populations of *E. cyclocarpum*. Our finding that suitable areas for the species are not expected to be severely impacted by the negative effects of climate change bodes well for its future use in tree planting projects. While a growing body of knowledge on the sexual and asexual propagation and management of *E. cyclocarpum* exists (Rodríguez-Sahagún et al. 2007; Laborde and Corrales-Ferrayola 2012; Ekamawanti et al. 2013; Obando and Moya 2013), limited attention has thus far been given

to the genetic quality of planting material. Ensuring that reproductive material of *E. cyclocarpum* used in tree planting efforts will survive, thrive, and establish viable and resilient populations in the long term requires (1) adequate planning to identify the seed sources that are best matched to the conditions of the planting site and (2) application of good seed collection protocols (Thomas et al. 2014, 2015a). To ensure suitability of planting material, identification and selection of germplasm should ideally be guided by the strength of the interaction between genotype performance and current and expected future environmental conditions (genotype-by-environment interactions), which are studied using multi-location progeny or provenance trials and climate modelling, respectively (Sgro et al. 2011; Breed et al. 2013). Empirical evidence has suggested that there is a significant effect of the location of origin of the seeds on early indicators of plant vigor in *E. cyclocarpum* (Rocha and Aguilar 2001). However, to the best of our knowledge, no progeny or provenance trials existed for *E. cyclocarpum* in Colombia, and at the time of this publication the first progeny trial was only in the initial phase of establishment by some of the authors. In the absence of multi-year observations from such trials, a combination of the neutral genetic characterization data and climate modelling we have reported here can be used as a best guess approach (Thomas et al. 2017). Although we did not find strong evidence for genetic differentiation between the populations we sampled, a precautionary approach would be to use the genetic clusters we identified based on correspondence analysis (Fig. 3) as proxies for seed transfer zones. This approach considers both the genetic differentiation we found between the southern Cauca River valley population of La Paila and all northern populations, and the possibility that the middle Cauca River valley and Caribbean region might each have given rise to different clusters at one point in prehistory, which were later mixed as a consequence of natural and human-influenced processes. As the impact of climate change on suitable areas of *E. cyclocarpum* is expected to be relatively low, reproductive material for tree planting activities in the geographical range of a given cluster should ideally be sourced from the same cluster range. This recommendation may later be adjusted based on observations from field trials or complementary genetic characterization studies.

For ensuring genetic diversity in planting stock, source populations should be large (ideally at least 500 reproductively mature individuals), and seeds should be obtained from a high number (30–60) of mother trees per population, among a series of other considerations recently summarized by Basey et al. (2015). Our results suggest that particularly the northern populations we sampled hold potential for sourcing good quality planting material. These should be priority sites for the in situ conservation of *E. cyclocarpum*. The fact that the trees we sampled in the buffer zone of Tayrona national park (ZAT) combined high genetic diversity with a low inbreeding coefficient and some degree of genetic differentiation with nearby populations, indicates that the park is likely to be effective in conserving the genetic diversity of local *E. cyclocarpum* populations. By contrast, reproductive material collected in the southern sites (COT–SFE, PIN, PAI) which scored high for the inbreeding coefficient should be mixed with materials from as many suitable populations as possible located within the range of a given cluster, while aiming for collection of seeds from a minimum of 60 mother trees. In some cases, particularly cluster 1 in whose range SDTF is highly degraded (Alvarado-Solano and Ospina 2015), it may be acceptable to mix reproductive material from either clusters 1 and 2, or 2 and 3, but not 1 and 3 (Fig. 3), to meet these targets. To promote the in situ persistence of the southern populations, they should be enriched with planting material collected from as many trees as possible in their respective cluster ranges.

Looking ahead, further research is needed on the effects of fragmentation on the reproductive success of *E. cyclocarpum* trees and the genetic profiles and vigor of progeny. To ensure availability of planting material of good genetic quality, a functional network of trained seed harvesters is needed. In the longer term, it would be desirable to establish seed orchards with selected elite trees in each of the different seed zones we have identified here.

**Acknowledgements** The authors wish to thank the Colombian companies Ecopetrol and Empresas Publicas de Medellin, the Government of the Colombian department of Antioquia, the CGIAR Fund Donors (<http://www.cgiar.org/who-we-are/cgiar-fund/fund-donors-2>) and the CGIAR research program on Forest Trees and Agroforestry for financial support. We are grateful to Jérôme Duminil and three anonymous reviewers for helpful comments on a previous version of this paper.

## References

- Aguilar R, Ashworth L, Galetto L, Aizen MA (2006) Plant reproductive susceptibility to habitat fragmentation: review and synthesis through a meta-analysis. *Ecol Lett* 9:968–980
- Aguilar R, Quesada M, Ashworth L et al (2008) Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol Ecol* 17:5177–5188
- Alvarado-Solano DP, Ospina JTO (2015) Distribución Espacial Del Bosque Seco Tropical En El Valle Del Cauca, Colombia. *Acta Biológica Colomb* 20:141–153
- Alzate-Marin AL, Guidugli MC, Soriani HH et al (2009) An efficient and rapid DNA minipreparation procedure suitable for PCR/SSR and RAPD analyses in tropical forest tree species. *Braz Arch Biol Technol* 52:1217–1224
- Banda-R K, Delgado-Salinas A, Dexter KG et al (2016) Plant diversity patterns in neotropical dry forests and their conservation implications. *Science* 353:1383–1387
- Basey AC, Fant JB, Kramer AT (2015) Producing native plant materials for restoration: 10 rules to collect and maintain genetic diversity. *Nativ Plants J* 16:37–53
- Braconnot P, Otto-bliesner B, Harrison S et al (2007) Results of PMIP2 coupled simulations of the Mid-Holocene and Last glacial maximum—Part 1: experiments and large-scale features. *Clim Past* 3:261–277
- Breed MF, Stead MG, Ottewell KM et al (2013) Which provenance and where? Seed sourcing strategies for revegetation in a changing environment. *Conserv Genet* 14:1–10
- Colpaert N, Cavers S, Bandou E et al (2005) Sampling tissue for DNA analysis of trees: trunk cambium as an alternative to canopy leaves. *Silvae Genet* 54:265–269
- de Abreu Moreira P, Brandão MM, de Araujo NH et al (2015) Genetic diversity and structure of the tree *Enterolobium contortisiliquum* (Fabaceae) associated with remnants of a seasonally dry tropical forest. *Flora Morphol Distrib Funct Ecol Plants* 210:40–46
- Dent EA, VonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361
- Deutsches klimarechenzentrum (DKRZ) (1992) The ECHAM3 atmospheric general circulation model. Technical Report 6. Hamburg
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13–15
- Duminil J, Mona S, Mardulyn P et al (2015) Late Pleistocene molecular dating of past population fragmentation and demographic changes in African rain forest tree species supports the forest refuge hypothesis. *J Biogeogr* 42:1443–1454
- Ekamawanti HA, Setiadi Y, Sopandie D, Santosa DA (2013) The Role of arbuscular mycorrhizal fungus (*Gigaspora margarita*) on mercury and nutrients accumulation by *Enterolobium cyclocarpum* Seedlings. *Microbiol Indones* 7:167–176
- Escalante EE (1985) Promising agroforestry systems in Venezuela. *Agrofor Syst* 3:209–221
- Etter A (2015) La transformación del uso de la tierra y los ecosistemas durante el período colonial en Colombia. In: Meisel Roca A, Ramírez GMT (eds) *La economía colonial de la Nueva Granada*, primera edn. FCE, Banco de la República, Bogotá, pp 62–103
- Etter A, van Wyngaarden W (2000) Patterns of landscape transformation in Colombia, with emphasis in the andean region. *AMBIO A J Hum Environ* 29:432–439



- Etter A, McAlpine C, Possingham H (2008) Historical patterns and drivers of landscape change in Colombia since 1500: a regionalized spatial approach. *Ann Assoc Am Geogr* 98:2–23
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491
- Frankel OH, Brown AHD, Burdon J (1995) The genetic diversity of wild plants. The conservation of plant biodiversity, First edit. University Press, Cambridge, pp 10–38
- Galluzzi G, Dufour D, Thomas E et al (2015) An integrated hypothesis on the domestication of *Bactris gasipa*. *PLoS ONE* 10:e0144644
- García H, Corzo G, Isaacs P, Etter A (2014) Distribución y estado actual de los remanentes del bioma: de bosque seco tropical en Colombia: insumos para su gestión. In: Pizano C, García H (eds) *El Bosque seco tropical en Colombia*. Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, Bogotá, pp 229–251
- Gonzales E, Hamrick JL, Smouse PE et al (2010) The impact of landscape disturbance on spatial genetic structure in the Guanacaste tree, *Enterolobium cyclocarpum* (Fabaceae). *J Hered* 101:133–143
- Hengl T, de Jesus JM, MacMillan R et al (2014) SoilGrids1 km—global soil information based on automated mapping. *PLoS ONE* 9:e105992
- Hijmans RJ (2012) Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* 93:679–688
- Hijmans RJ, Cameron SE, Parra JL et al (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25:1965–1978
- Janzen DH (1981) *Enterolobium cyclocarpum* seed passage rate and survival in horses, Costa Rican Pleistocene seed dispersal agents. *Ecology* 62:593–601
- Janzen DH (1982) Variation in average seed size and fruit seediness in a fruit crop of a Guanacaste tree (Leguminosae: *Enterolobium cyclocarpum*) (Costa Rica). *Am J Bot* 69:1169–1178
- Janzen DH (1983) *Costa Rican natural history*. University of Chicago Press, Chicago
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405
- Jombart T, Devillard S, Dufour A, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* 101:92–103
- Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281
- Kindt R, Coe R (2005) *Tree diversity analysis. A manual and software for common statistical methods for ecological biodiversity studies*. World Agroforestry Centre (ICRAF), Nairobi
- Laborde JL, Corrales-Ferrayola I (2012) Direct seeding of *Brosimum alicastrum* SW. (Moraceae) and *Enterolobium cyclocarpum* Griseb. (Mimosaceae) in different habitats in the dry tropics of central Veracruz. *Acta Bot Mex* 100:107–134
- Lagemann J, Heuvelink J (1983) Characterization and evaluation of agroforestry systems: the case of Acosta-Puriscal, Costa Rica. *Agrofor Syst* 1:101–115
- Lowe AJ, Boshier D, Ward M et al (2005) Genetic resource impacts of habitat loss and degradation: reconciling empirical evidence and predicted theory for neotropical trees. *Heredity* 95:255–273
- Mahecha L (2002) El silvopastoreo: una alternativa de producción que disminuye el impacto ambiental de la ganadería bovina. *Rev Colomb Ciencias Pecu* 15:226–231
- Marchant R, Behling H, Berrío JC et al (2001) Mid- to Late-Holocene pollen-based biome reconstructions for Colombia. *Quat Sci Rev* 20:1289–1308
- Marchant R, Behling H, Berrío JC et al (2002) Pollen-based biome reconstructions for Colombia at 3000, 6000, 9000, 12000, 15000, and 18000 14C yr ago: late quaternary tropical vegetation dynamics. *J Quat Sci* 17:113–129
- Miles L, Newton AC, DeFries RS et al (2006) A global overview of the conservation status of tropical dry forests. *J Biogeogr* 33:491–505
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci* 70:3321–3323
- Novaes RM, Rodrigues JG, Lovato MB (2009) An efficient protocol for tissue sampling and DNA isolation from the stem bark of Leguminosae trees. *Genet Mol Res* 8:86–96
- Obando MF, Moya R (2013) Silviculture conditions and wood properties of *Samanea saman* and *Enterolobium cyclocarpum* in 19-year-old mixed plantations. *For Syst* 22:58–70
- Pennington RT, Prado DE, Pendry C (2000) Neotropical seasonally dry forests and Quaternary vegetation changes. *J Biogeogr* 27:261–273

- Peters MB, Hagen C, Dorset W et al (2008) Isolation and characterization of microsatellite loci in the Guanacaste tree, *Enterolobium cyclocarpum*. *Mol Ecol Resour* 8:129–131
- Petit RJ, Aguinagalde I, Beaulieu JL et al (2003) Glacial refugia hotspots but not melting pots of genetic diversity. *Science* 80(300):1563–1565
- Pizano C, González-M R, Gonzáles MF et al (2014) Las plantas de los bosques secos de Colombia. In: Pizano C, García H (eds) *El Bosque seco tropical en Colombia*. Instituto de Investigación de recursos Biológicos Alexander von Humboldt, Bogota, pp 49–93
- Prado D, Gibbs P (1993) Patterns of species distributions in the dry seasonal forests of South America. *Ann Missouri Bot Gard* 80:902–927
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Ramírez Villegas J, Jarvis A (2010) Downscaling global circulation model outputs: the delta method. International Center for Tropical Agriculture (CIAT), Cali
- Reinoso-Pérez M (2014) Agroforestry: a viable alternative for sustainable agricultural production. *Sci Agric* 1:32–39
- Rocas NA (2002) *Enterolobium cyclocarpum*. In: Vozzo JA (ed) *Tropical tree seed manual*. United States Department of Agriculture Forest Service, USA, pp 449–451
- Rocha O, Aguilar G (2001) Variation in the breeding behavior of the dry forest tree *Enterolobium cyclocarpum* (Guanacaste) in Costa Rica. *Am J Bot* 88:1600–1606
- Rocha OJ, Lobo JA (1996) Genetic variation and differentiation among five populations of the Guanacaste tree (*Enterolobium cyclocarpum* Jacq.) in Costa Rica. *Int J Plant Sci* 157:234–239
- Rodríguez-Sahagún A, Castellanos-Hernandez OA, Acevedo-Hernandez GJ, Excised A (2007) In vitro propagation of *Enterolobium cyclocarpum* (guanacaste) from nodal explants of axenic seedlings. *E-Gnosis* 5:1–14
- Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nat Biotechnol* 18:233–234
- Sgro CM, Lowe AJ, Hoffmann AA (2011) Building evolutionary resilience for conserving biodiversity under climate change. *Evol Appl* 4:326–337
- Thomas E, van Zonneveld M, Loo J et al (2012) Present Spatial Diversity Patterns of *Theobroma cacao* L. in the Neotropics Reflect Genetic Differentiation in Pleistocene Refugia Followed by Human-Influenced Dispersal. *PLoS ONE* 7:e47676
- Thomas E, Jalonen R, Loo J et al (2014) Genetic considerations in ecosystem restoration using native tree species. *For Ecol Manag* 333:66–75
- Thomas E, Jalonen R, Loo J, Bozzano M (2015a) Avoiding failure in forest restoration: the importance of genetically diverse and site-matched germplasm. *Unasylva* 66:29–36
- Thomas E, Alcázar Caicedo C, McMichael CH et al (2015b) Uncovering spatial patterns in the natural and human history of Brazil nut (*Bertholletia excelsa*) across the Amazon Basin. *J Biogeogr* 42:1367–1382
- Thomas E, Alcazar C, Moscoso LG et al (2017) The importance of species selection and seed sourcing in forest restoration for enhancing adaptive potential to climate change: Colombian tropical dry forest as a model. In: *CBD Technical Series: Biodiversity and Climate Change*. CBD secretariat, Montréal
- Vina A, Cavelier J (1999) Deforestation Rates (1938–1988) of tropical lowland forests on the andean foothills of Colombia. *Biotropica* 31:31–36
- Werneck FP, Costa GC, Colli GR et al (2011) Revisiting the historical distribution of seasonally dry tropical forests: new insights based on palaeodistribution modelling and palynological evidence. *Glob Ecol Biogeogr* 20:272–288
- Werneck FP, Nogueira C, Colli GR et al (2012) Climatic stability in the Brazilian Cerrado: implications for biogeographical connections of South American savannas, species richness and conservation in a biodiversity hotspot. *J Biogeogr* 39:1695–1706
- Young AG, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends Ecol Evol* 11:413–418