Pleistocene climatic fluctuations promoted alternative evolutionary histories in *Phytelephas aequatorialis*, an endemic palm from western Ecuador

Sebastián Escobar1 | Andrew J. Helmstetter2 | Scott Jarvie1,3 | Rommel Montúfar4 | Henrik Balslev1 | Thomas L.P. Couvreur2,4

1Section for Ecoinformatics and Biodiversity, Department of Biology, Aarhus University, Aarhus C, Denmark
2IRD, UMR DIADE, Université de Montpellier, Montpellier, France
3Center for Biodiversity Dynamics in a Changing World (BIOCHANGE), Department of Biology, Aarhus University, Aarhus C, Denmark
4Facultad de Ciencias Exactas y Naturales, Pontificia Universidad Católica del Ecuador, Quito, Ecuador

Correspondence
Sebastián Escobar, Section for Ecoinformatics and Biodiversity, Department of Biology, Aarhus University, DK 8000 Aarhus C, Denmark.
Email: sescobarvasquez@gmail.com

Funding information
SENESCYT; Independent Research Fund Denmark, Grant/Award Number: 9040-00136B; Agence Nationale de la Recherche, Grant/Award Number: ANR-15-CE02-0002-01; International Mixed Laboratory BIO_INCA; International Palm Society

Handling editor: Carina Hoorn

Abstract

**Aim:** Pleistocene (2.58 Ma–11.7 ka) climatic fluctuations have shaped intraspecific genetic patterns worldwide; however, their impact on species in many regions remains unknown. In order to determine the impact of Pleistocene climatic fluctuations on the tropical rain forests of western Ecuador, we explored the evolutionary history of the endemic palm *Phytelephas aequatorialis*.

**Location:** Western Ecuador, north-western South America.

**Taxon:** *Phytelephas aequatorialis* (Arecaceae).

**Methods:** One hundred and seventy-six nuclear genes were sequenced in 91 individuals for phylogenomic and population structure analyses. The time of divergence between identified genetic lineages was estimated using a coalescent phylogenomic analysis. Palaeoecological niche modelling analyses were performed to determine areas of historical climatic suitability since the Last Glacial Maximum (LGM; 22 ka) that potentially acted as forest refugia during the Pleistocene. A Wilcoxon test and Pearson correlations were used to explore how current levels of genetic diversity, in terms of expected heterozygosity ($H_s$), have been shaped by several palaeoclimatic and geographic factors.

**Results:** Phylogenomic and population structure analyses revealed two main genetic lineages with a north–south distribution, which diverged 1.14 Ma during the Pleistocene. Two potential Pleistocene refugia were identified, one along the Pacific coast of Ecuador and one in the Andean foothills of south-western Ecuador. The location of these refugia agrees with the spatial location of the two genetic lineages. Within the Andean foothills, $H_s$ was lower for the southern lineage than for the northern lineage. $H_s$ significantly increased with decreasing latitude across the species as a whole.

**Main conclusions:** Pleistocene climatic fluctuations promoted intraspecific divergence in *P. aequatorialis* within the rain forests of western Ecuador. The Andean foothills of south-western Ecuador could be an important area for rain forest evolution.
Climatic fluctuations during the Pleistocene (2.58 Ma – 11.7 ka) have shaped the current genetic patterns of plant and animal species across species-rich tropical rain forests worldwide. The forest refugium hypothesis stipulates that Pleistocene glacial cycles promoted retraction and fragmentation of tropical rain forests, isolating populations within climatically suitable areas enhancing allopatric speciation (Haffer, 1969). Even though the role of the Pleistocene in promoting speciation across the Neotropics has been questioned (Bush, 1994; Colinaux et al., 1996, 2000; Hoorn et al., 2010; Leite et al., 2016), it has played a major role in shaping intraspecific genetic diversity (Hewitt, 2000). Intraspecific genetic lineages have originated in areas where tropical species took refuge during the Pleistocene because isolated populations restricted gene flow (Cabanne et al., 2016; Carnaval & Bates, 2007; Carnaval et al., 2009; Diniz-Filho et al., 2016; Thomas et al., 2012; Vasconcellos et al., 2019). Additionally, higher levels of genetic diversity are usually observed in palaeoclimatically suitable areas where tropical species persisted during the Pleistocene (Carnaval et al., 2009; Faye et al., 2016; Mousalli et al., 2009; Walker et al., 2009) because long-term populations tend to accumulate unique genetic diversity when compared to more recently expanding ones (Excoffier et al., 2009). Within the Neotropics, the intraspecific genetic implications of the Pleistocene have not been explored west of the Andes, hampering our understanding of the recent evolutionary history of this region.

Western Ecuador is a species-rich and heterogeneous region located west of the Ecuadorian Andes. The rain forests that occur in western Ecuador form a highly diverse ecoregion in a relatively small area of 34,000 km² (Gentry, 1986; Olson et al., 2001). For instance, 870 native plant species have been recorded in just 1 km² with −20% of endemicity (Dodson & Gentry, 1978). Unfortunately, this biodiversity is highly threatened due to widespread deforestation and land conversion for agricultural purposes (Cerón et al., 1999; Dodson & Gentry, 1991; Ferrer-Paris et al., 2019; Olson et al., 2001; Sierra, 2013). The rain forests of western Ecuador form a limit between the humid Chocó to the north and the dry Tumbes region to the south. Therefore, these rain forests present a north/south gradient of annual precipitation, averaging 7,000 mm to the north and decreasing to 2,000 mm to the south (Dodson & Gentry, 1991). In addition to climatic variation, the rain forests of western Ecuador present a wide altitudinal range from sea level up to 1,500 m.a.s.l. They are distributed in the lowlands and along the small cordilleras of the Pacific coast, as well as along the adjacent western Andean foothills that gradually narrow towards the south (Cerón et al., 1999; Dodson & Gentry, 1991; Sierra, 2013). Thus, the rain forests of western Ecuador could have responded differently to Pleistocene climatic fluctuations because of their north/south topographic and climatic heterogeneity.

The location of rain forest refugia during the Pleistocene in western Ecuador is still poorly understood. Overall, Neotropical rain forests are thought to have retracted leading to isolated forest patches during glacial cycles where temperatures dropped −5°C compared to pre-industrial values (Heine, 2000; Hewitt, 2000). Based on current patterns of plant distribution, lowland rain forests are hypothesized to have survived the Pleistocene within the Chocó refugium, situated along the Pacific coast in north-western Ecuador and western Colombia (Gentry, 1982; Prance, 1982). Although evidence of the climatic stability of this region has only been documented for the Holocene (10 ka – present; Behling et al., 1998), the high levels of precipitation linked to its proximity to the Pacific Ocean may have allowed rain forests to persist in these lowlands during the Pleistocene (Hooghiemstra & Van Der Hammen, 1998).

In addition, the rain or cloud forests that occur on the Andean foothills of western Ecuador (hereinafter referred as premontane rain forests) could have also acted as refugia during the Pleistocene. Glacial cooling promoted the down-slope displacement of Neotropical premontane and montane rain forests (Colinaux et al., 1996; Valencia et al., 2010). However, the moisture of the adjacent lowlands could have determined whether they were able to displace and spread into the lowlands (‘moist forest model’; Ramirez-Barahona & Eguiarte, 2013) or whether they contracted into refugia at mid-elevations by the opposing forces of aridity and cooling (‘dry refugia model’; Ramirez-Barahona & Eguiarte, 2013). The ‘moist forest model’ predicts population connectivity in the lowlands, diffuse population genetic structure and high genetic diversity because of spatial heterogeneity (Ramirez-Barahona & Eguiarte, 2013). In this sense, the Andean foothills of north-western Ecuador may have not been explored west of the Andes, hampering our understanding of the recent evolutionary history of this region.
acted as Pleistocene refugia because premontane rain forests potentially moved down-slope into the Pacific coast and remained in the Chocó refugium (Hooghiemstra & Van Der Hammen, 1998). The ‘dry refugia model’, in contrast, predicts the isolation of populations at mid-altitude with subsequent genetic structuring and loss of genetic diversity because of altitudinal contraction and population reduction (Ramírez-Barahona & Eguiarte, 2013). Given that the lowlands of south-western Ecuador remained dry during recent glacial cycles (Heusser & Shackleton, 1994), the adjacent Andean foothills could have harboured premontane rain forests during the Pleistocene because they were not able to displace into the arid lowlands.

Here we explore how Pleistocene climatic fluctuations have shaped intraspecific genetic diversity patterns of an endemic plant species from the rain forests of western Ecuador. We reconstruct the evolutionary history of the vegetable ivory palm *Phytelephas aequatorialis* Spruce, by inferring its intraspecific phylogenetic relationships, population structure, and palaeoclimatic niche distribution. This species is an emblematic element of these forests (Acosta-Solis, 1948; Barford, 1991; Cerón et al., 1999), occurring from sea level to 1,500 m.a.s.l. (Borchsenius et al., 1998). According to previous molecular dating studies, *P. aequatorialis* diverged from its sister species *Phytelephas tumacana* O.F. Cook approximately 3 – 4 Ma, shortly before the Pleistocene (Barford et al., 2010; Escobar et al., 2020; Trénel et al., 2007). Speciation in the genus *Phytelephas* apparently occurred by the action of Andean uplift (Escobar et al., 2020) followed by shifts in pollination mechanisms and vegetative morphology (Barford et al., 2010). Thus, posterior intraspecific diversification in *Phytelephas* species may have occurred during the Pleistocene. In this context, *P. aequatorialis* appears as an adequate model species for studying how Pleistocene climatic fluctuations impacted the evolution of rain forest species in western Ecuador. We used a genomic approach by targeting 176 nuclear genes using a custom palm baiting kit (Heyduk et al., 2016) sequenced for 91 adult individuals identified in the field as *P. aequatorialis* from 46 different localities across western Ecuador (Table S1). In each locality, we sampled between one and three individuals, which were separated by at least 100 m between them to avoid the sampling of closely related individuals. Two previously collected individuals of the sister species *P. tumacana* from south-western Colombia were used as outgroup in the phylogenomic analyses. We collected tissue from the youngest leaf of the individuals, which was then silica gel-dried in plastic bags. We isolated total genomic DNA from each sample using the MATAB and chloroform separation methods as described in Mariac et al. (2014), and checked DNA quality and quantity using a NanoDrop™ One spectrophotometer (Thermo Scientific™).

### 2.2 Library preparation, sequencing, filtering and assembly

Laboratory work was performed at the facilities of the Institut de Recherche pour le Développement (IRD) in Montpellier, France. We prepared genomic libraries using a modified protocol from Rohland and Reich (2012) and captured nuclear markers using a bait kit that targets 176 palm-specific nuclear genes (Heyduk et al., 2016). Total DNA for each individual was sheared to a mean target size of 350 bp using a Bioruptor® Pico sonication device (Diagenode). We then replica, ligated and nick filled-in the DNA before amplifying it for 10 cycles as a pre-hybridization step. Next, after cleaning and quantifying the DNA, we bulked the samples in two libraries and added the biotin-labelled baits to the libraries in order to hybridize the targeted regions of the bait kit used (Heyduk et al., 2016). We used streptavidin-coated magnetic beads to immobilize the hybridized biotin-labelled baits applying a magnetic field, and then discarded the unbound DNA contained in the supernatant. We amplified the enriched DNA fragments eluted from the beads in a RT-PCR during 15 cycles to complete adapters. Final libraries were sequenced with an Illumina® MiSeq (paired-end, length 150 bp) at the facilities of the CIRAD (Montpellier, France) using two lanes.

High-performance computing analyses were done at the South Green Platform at the IRD. We followed Couvreur et al. (2019) for cleaning and filtering the reads. We controlled the quality using FastQC ([http://www.bioinformatics.babraham.ac.uk/projects/fastqc](http://www.bioinformatics.babraham.ac.uk/projects/fastqc)), demultiplexed using the ‘demultadapt’ script ([https://github.com/Maillo/demultadapt](https://github.com/Maillo/demultadapt)) with a 0-mismatch threshold and removed the adapters using Cutadapt v.1.2.1 (Martin, 2011). Reads with length <35 bp and quality mean values (Q) <30 were discarded using a custom script for quality filtering ([https://github.com/SouthGreenPlatform/arcad-hts/blob/master/scripts/arcad_hts_2_Filter_Fastq_On_Mean_Quality.pl](https://github.com/SouthGreenPlatform/arcad-hts/blob/master/scripts/arcad_hts_2_Filter_Fastq_On_Mean_Quality.pl)). Then, we paired forward and reverse sequences using a comparison script adapted from TOGGLE (Tranchant-Dubreuil et al., 2018), and checked the quality again using FASTQC. The last 6 bp of the reverse sequences were trimmed to ensure the removal of barcodes in sequences <150 bp using the ‘fastx’ script ([https://github.com/agordon/fastx_toolkit](https://github.com/agordon/fastx_toolkit)).

### 2 MATERIALS AND METHODS

#### 2.1 Geographic sampling and DNA extraction

We sampled and georeferenced 91 adult individuals identified in the field as *P. aequatorialis* from 46 different localities across western Ecuador.
2.3 Phylogenomic inference

DNA sequences were processed using HybPiper v.1.2 (Johnson et al., 2016) to identify target exonic regions and off-target introns, which are usually more variable than exons. As a result, we obtained ‘supercontigs’ which are formed by targeted regions and surrounding off-target sequences. We filtered the dataset by keeping only the loci with >75% of their length reconstructed in >75% of the individuals. The remaining supercontigs were aligned using the ‘-auto’ option in MAFFT v.7.305 (Katoh & Standley, 2013) and were cleaned using Gblocks v.0.91b (Castresana, 2000) with default parameters and all allowed gap positions.

HybPiper also detects potential paralogous loci, which were verified by building exon trees in RAxML v.8.2.9 (Stamatakis, 2014), including all putative paralogs. We used a GTR + G substitution model because it is the most common and general model for DNA analyses (Stamatakis, 2016), producing similar inferences as other models (Abadi et al., 2019). We also used the option ‘-f a’ to perform 100 rapid bootstraps as node support. For this purpose, we used a reference where exons are split instead of being merged into genes. We examined the exon trees manually to check if the potential paralog sequences that were flagged by HybPiper clustered together, which would mean that the potential paralogs are more closely related to potential paralogs in other individuals than to the alternative sequences in the same individual (Helmstetter et al., 2020). In this case, we treated the exon as a paralog. If >50% of exons belonging to a certain gene were determined as paralogs, then the gene was marked as ‘true’ paralog and removed from further steps.

We used the resulting alignments to infer the phylogenomic relationships within P. aequatorialis using the two individuals of P. tumacana as outgroup. We used two approaches to construct the intraspecific tree, coalescence and maximum likelihood. For the coalescence approach, we used the alignments to infer individual gene trees in RAxML with a GTR + G model, 100 rapid bootstraps and no partitions. We then used these gene trees to construct a coalescent tree in ASTRAL-III v.5.6.3 (Zhang et al., 2018), collapsing branches with support <10 and annotating the local posterior probability (pp) as node support. For the maximum likelihood approach, we first added empty sequences where individuals were missing in our alignments and concatenated them with phyx (Brown et al., 2017), generating one partition per loci. We built the maximum likelihood tree in RAxML using a GTR + G model and 100 rapid bootstraps. Phylogenomic trees were visualized in FigTree v.1.4 (http://tree.bio.ed.ac.uk/software/figtree/). Three individuals of P. aequatorialis showed high node support with P. tumacana in the two intraspecific trees, suggesting close phylogenomic relatedness between these individuals. Therefore, we used this mixed clade to root our trees and excluded these three individuals of P. aequatorialis from further analyses (see Results).

2.4 Population structure

We studied the population structure of P. aequatorialis by determining the number of genetic clusters within its distribution. We processed the DNA sequences of 88 individuals with the pipeline SECAPR v.1.14 (Andermann et al., 2018). This pipeline is useful when SNPs need to be called because it creates a pseudo reference formed by the consensus sequences of the target loci, with the exception of paralog loci that were identified and automatically removed during the process. The use of the aforementioned pseudo reference is more efficient than the use of the bait kit reference in the phylogenomic analyses because the latter is made from distantly related palm species (Helmstetter et al., 2020). We used BWA v.0.7.12 (Li & Durbin, 2009) to map this pseudo reference to our paired and cleaned reads. GATK v.4 (McKenna et al., 2010) was used to remove duplicates and to call SNPs. We filtered the data by keeping only biallelic SNPs with a mapping quality >40%, depth >25, and quality by depth >2 using BCFtools (Li, 2011). Additionally, we removed SNPs with allele frequency <0.01 and excluded monomorphic sites.

The number of clusters were determined with a Discriminant Analysis of Principal Components (DAPC; Jombart et al., 2010) using the package ‘adegenet’ (Jombart, 2008) in R v.3.5.3 (R Core Team, 2019). We used successive K-means with 100,000 iterations per value of K (number of clusters) up to K = 20. We examined the change in the Bayesian Information Criterion (BIC) of K = 1–20 to determine the number of clusters within P. aequatorialis. We then did a cross-validation of the DAPC analysis to test that the number of principal components chosen was correct.

2.5 Divergence time estimates

We estimated the divergence time between the resulting genetic lineages by constructing a dated coalescent tree using five random individuals from each clade/genetic cluster, along with the two individuals of P. tumacana as outgroup. We calculated the root-to-tip variance of their respective RAxML gene trees and rooted them with Phyx. We then calculated the variance using SortaDate (https://github.com/FePhyFoFum/SortaDate/) and chose the 32 most clock-like loci (most suitable for dating) to reduce the highly demanding computational requirements of further steps (Smith et al., 2018). The best nucleotide substitution model of evolution was determined for each alignment using ModelTest (https://github.com/ddarriba/modeltest). The corresponding 32 fasta alignments were converted to nexus format using PGDSpider v.2.1.1.5 (Lischer & Excoffier, 2012).

The coalescence analysis was run using Bayesian inference in BEAST v.2.5.2 (Bouckaert et al., 2019). We used the template ‘StarBeastWithSTACEYops’ and assigned the individuals to their respective lineage. We used a strict clock of 0.001 with the ‘estimate’ box checked and the best nucleotide substitution model of evolution for each locus as determined in ModelTest. The tree was generated with a Coalescent Constant Population prior because it is particularly suitable for describing intraspecific relationships (Ritchie et al., 2017). The tree was constrained with a secondary calibration placed at the stem node between P. aequatorialis and its sister
species *P. tumacana*. We used a uniform prior and constrained the age of the node between the 95% highest posterior density heights (95% HPD) of the speciation time between *P. aequatorialis* and *P. tumacana* (see Escobar et al., 2020 for details), which were used as the minimum and maximum values of the prior (3.01–3.76 Ma). We ran three independent Markov Chain Monte Carlo (MCMC) analyses with 100 million generations sampling every 10,000 trees in BEAST. We checked that the MCMC runs reached convergence using Tracer v.1.7 (Rambaut et al., 2018), ensuring that all effective sample sizes (ESS) >200. The log and tree files of the three runs were combined with LogCombiner v.2.5.2 (Bouckaert et al., 2019) with a burn-in of 10%. TreeAnnotator v.2.5.2 (Bouckaert et al., 2019) was used to generate a maximum clade credibility (MCC) tree by summarizing the remaining 27,003 trees, and to calculate 95% HPD. The tree was visualized in FigTree.

### 2.6 Palaeoecological niche modelling

We modelled the ecological niche of the species under current conditions, and then projected it to the LGM to determine the potential effect of Pleistocene glaciations on the distribution of *P. aequatorialis*. In addition to occurrence records collected from our fieldwork, we downloaded records from four online repositories: the Global Biodiversity Information Facility (GBIF.org, 2020), the Berkeley Ecoengine (Ecoengine), the Botanical Information and Ecological Network (BIEN v.4.1) and the Latin American Seasonality Dry Tropical Forest Floristic Network (DRYFLOR). We accessed occurrence records from the repositories using the R-packages ‘BIEN’ (Maitner et al., 2018) and ‘spocc’ (Chamberlain & Boettiger, 2017), and direct download (http://www.dryflor.info/). We cleaned the downloaded occurrence data before use with the package ‘CoordinateCleaner’ (Zizka et al., 2019), removing records that met any of the following criteria: recorded prior to 1950; duplicated records; coordinate uncertainty greater than 5 km; impossible or unlikely coordinate; within 10 km of a country capital or political unit centroids; and within 100 m of biodiversity institutions. All occurrence records were spatially thinned at a distance of at least 10 km to reduce spatial sampling bias using the package ‘spThin’ (Aiello-Lammens et al., 2015). We retained and used a total of 86 records after cleaning and filtration of data, which were transformed to a Behrmann projection to be used along with resampled climatic variables.

The current (average 1950–2000) and LGM (22 ka) variables were downloaded at a spatial resolution of 2.5 arc-minute from WorldClim 1.4 (https://worldclim.org; Hijmans et al., 2005) and ENVIREM (http://envirem.github.io; Title & Bemmels, 2018). Variables were then resampled to 5 km resolution using bilinear interpolation with a Behrmann projection. Climatic variables from the LGM were downloaded for the three General Circulation Models CCSM4 (Gent et al., 2011), MIROC-ESM (Watanabe et al., 2011) and MPI-ESM-P (Giorgetta et al., 2013). To minimize potential issues with collinearity between predictor variables, we ensured the Pearson’s R correlation was <0.7 (Dormann et al., 2013). We identified variables that could potentially limit the distribution of the species. Potential evapotranspiration (PET) is one of the strongest predictors of plant diversity patterns at broad scales (Eiserhardt et al., 2011 and references therein). More specifically, low precipitation and low temperatures can constrain the distribution of palms (Eiserhardt et al., 2011). We therefore chose PET of the wettest quarter and PET of the driest quarter from ENVIREM, and precipitation of the driest month (BIO 14) and temperature annual range (BIO 7) from WorldClim as our predictor variables. We did not choose minimum temperature of the coldest month (BIO 6) as a predictor variable because it was highly correlated with PET of the driest quarter (>0.7).

To model the potential current and LGM distributions of *P. aequatorialis*, we used the new ‘maxnet’ R-package to fit a maximum entropy (Maxent) model (Phillips et al., 2017). To reduce the effects of sampling bias with this presence-only algorithm, we used a weighted-target group approach (Phillips et al., 2009). This approach uses a background sample likely to contain the same sampling bias as the occurrence records for the focal species, allowing the bias to be cancelled out of the fitting and thus more accurate presence-only models can be produced (Baumgartner et al., 2018; Merow et al., 2017; Phillips et al., 2009). For this purpose, we used 1,149 occurrence records of palm species within a 200 km buffer around the records of *P. aequatorialis*. The palm records were downloaded from the same sources listed above and cleaned in the same way. The buffer size selected includes environments that are probably accessible to the species given its dispersal limitations and the configuration of natural barriers (Barve et al., 2011). We used the linear, quadratic and product features in the Maxent model to produce smoother response curves that reduce overfitting, which is especially relevant when projecting to novel conditions (Merow et al., 2013). The performance of the model was assessed by calculating the average Continuous Boyce Index (CBI; Hirzel et al., 2006) through a fivefold systematic spatial blocking technique using the package ‘blockCV’ (Valavi et al., 2019). CBI values range between ~1 and 1, where values >0 indicate the model’s output is positively correlated with the true probability of presence and values <0 indicate it is negatively correlated with the true probability of presence (Fielding & Bell, 1997). Model fit was inspected by the average omission rate based on the minimum training presence value (OR$_{\text{MTP}}$). OR$_{\text{MTP}}$ values range from 0 for models that are not overfit to 1 for models that are overfit. For the LGM, we averaged the projections of the three General Circulation Models continuous range maps to form a consensus continuous map. Current and LGM continuous suitability maps were converted to binary maps indicating suitable/unsuitable habitat. For this purpose, we used a threshold that maximizes the sum of sensitivity and specificity, a frequently recommended approach that tends to reflect the prevalence of the modelled species well (Liu et al., 2013, 2016). We then overlaid the two binary maps to show areas of historical climatic suitability that potentially worked as refugia during the LGM. We used this overlaid map as an approximation of the potential refugia location during previous Pleistocene glacial cycles.
2.7 Determinants of genetic diversity

We explored how palaeoclimatic and geographic factors may have contributed to shape current levels of genetic diversity in *P. aequatorialis*. For each individual, we obtained the expected heterozygosity (Hs) using the R-package ‘adegenet’ (Jombart, 2008) and the index of palaeoclimatic suitability during the LGM from the consensus continuous map using the package ‘raster’ (Hijmans et al., 2020). We first explored how the ‘moist forest’ and ‘dry refugia’ models have potentially shaped genetic diversity in the Andean foothills. To do so, we tested whether individuals from the Andean foothills that acted as Pleistocene refugia present lower Hs than the individuals from the foothills that did not act as refugia using a Wilcoxon rank sum test in R. We also performed Pearson’s R correlations to regress Hs against values of palaeoclimatic suitability during the LGM, latitude and longitude for all individuals. Additionally, genetic differentiation (FST) between the clusters was calculated with the package ‘hierfstat’ (Goudet, 2005) using the Nei’s method (Thomas et al., 2012).

3 RESULTS

3.1 Phylogenomic inference

We obtained a total of 47,167 million reads from the 91 adult individuals of *P. aequatorialis*. We identified 157 loci with >75% of its length reconstructed in >75% of individuals. Sixteen loci were considered as paralogs and therefore removed, resulting in 141 supercontigs. The resulting ASTRAL coalescent tree separated most of the *P. aequatorialis* individuals from the two individuals of *P. tumacana* (Figure 1a). Three individuals sampled in the northern part of the distribution of *P. aequatorialis* formed a clade with *P. tumacana* with high local posterior probability as node support (pp = 1). The other individuals of *P. aequatorialis* separated in two major clades, suggesting the presence of two distinct evolutionary units within the species. One of the clades was formed by 79 individuals sampled from northern and central populations (northern clade), whereas the other clade was formed by nine individuals sampled from southern populations (southern clade). Only the southern clade showed high node support.

FIGURE 1 Intraspecific genetic lineages of the palm *Phytelephas aequatorialis* in the rain forests of western Ecuador. (a) ASTRAL coalescent tree showing phylogenomic relationships between individuals with local posterior probabilities (pp) as node support. (b) Spatial distribution of the genetic clusters at K = 2 as determined by the DAPC analysis using 88 individuals, and of the individuals used as outgroup in the phylogenomic analyses. The shading represents the Andes and smaller cordilleras in the Pacific coast. The Chimbo River is represented by the light blue line. (c) Membership probability of the 88 individuals at K = 2 as determined by the DAPC analysis [Colour figure can be viewed at wileyonlinelibrary.com]
The RAxML maximum likelihood tree recovered the same two major clades in *P. aequatorialis* (Figure S1). Nevertheless, one individual of *P. aequatorialis* from the northern part of its distribution resolved as sister of the two major clades. The same three individuals of *P. aequatorialis* from the north formed a clade with *P. tumacana* with high bootstrap as node support (93). Again, only the southern clade showed high node support (bootstrap = 100; Figure S1).

### 3.2 Population structure and divergence time estimates

A total of 1,394 SNPs were identified from 118 different loci with 0% of missing data. We kept 50 principal components after cross-validation, as it was showed appropriate for inferring population structure in DAPC analyses (Figure S2). Two genetic clusters were determined within *P. aequatorialis* (Figure 1b) based on the values of BIC and on the analyses of principal components (Figures S3, S4). All individuals showed a membership probability to their corresponding cluster >0.9 (Figure 1c). The two resulting clusters are composed by the same individuals forming the two major clades in the ASTRAL coalescent tree (Figure 1a). The division between the two genetic lineages lies around −2.4° in latitude. We estimated that the two genetic clusters diverged during the Pleistocene 1.14 Ma (95% HPD: 1.01 – 2.48; Figure 2).

### 3.3 Palaeoeocological niche modelling

The mean CBI value was 0.91, indicating that the model performance is better than a random model and therefore adequate. The model did not over fit as indicated by the low ORMTP value of 0.02. The model accurately predicted the current known distribution of the species with exception of the northernmost part of the range (Figure 3a). The lowlands of north-western Ecuador appeared unsuitable even though *P. aequatorialis* has been reported in that area, whereas an area in the Andean foothills of Colombia, where the species has not been reported from, appeared as suitable. Overall, climatic suitability of *P. aequatorialis* decreased during the LGM and its potential distribution remained in the Pacific coast while decreasing on the Andean foothills (Figure 3b). The overlaid map formed by the two suitable/unsuitable maps for current and past conditions during the LGM showed two areas of historical climatic suitability (Figure 3c). The first one is located in the Pacific coast while a second smaller area lies on the Andean foothills of south-western Ecuador. The location of the larger area of historical climatic suitability agrees with the spatial distribution of the northern clade/genetic cluster, whereas the smaller area agrees with the southern clade/genetic cluster.

### 3.4 Determinants of genetic diversity

The nine individuals from the Andean foothills of south-western Ecuador that belong to the southern lineage (Figure 1; Table S1) showed lower *Hs* than the 19 individuals from the Andean foothills of north-western Ecuador that belong to the northern lineage (Wilcoxon rank sum test: *W* = 11, *p* = 0.0000278; Figure 4a). We did not detect a significant correlation between *Hs* and the index of palaeoclimatic suitability during the LGM (*R*² = 0.0019, *p* = 0.6854, *t* = −0.4064, *df* = 86; Figure 4b). A significant but weak increase in *Hs* with decreasing latitude was inferred (*R*² = 0.099, *p* = 0.0028, *t* = 3.0753, *df* = 86; Figure 4c). *Hs* was not correlated with longitude (*R*² = 0.0115, *p* = 0.3183, *t* = 1.0037, *df* = 86; Figure 4d). Genetic differentiation (*F*ₜₛ) between the two genetic clusters was 0.023.

### 4 DISCUSSION

Our phylogenomic and population structure analyses revealed the presence of two distinct latitudinally oriented evolutionary lineages in the palm species *Phytelephas aequatorialis* within the rain forests of western Ecuador (Figure 1). Although our intraspecific trees were in general poorly supported, the southern clade showed high support under the two phylogenomic approaches used (Figure 1b, 1c), adding more evidence to their distinct evolutionary history. A similar pattern of phylogeographic structure was detected for the widely distributed Neotropical palm *Oenocarpus bataua*, in which two genetic clusters...
were also identified in western Ecuador: one located in north-western Ecuador/south-western Colombia, and a second one in south-western Ecuador and extending to the east of the Andes (Escobar et al., 2018). Although the genetic structure patterns of both species are not completely identical due to the cross-Andean distribution of O. bataua, they nevertheless suggest the existence of a north–south phylogeographic break in south-western Ecuador. The Chimbo River has been proposed as a geographic barrier for some reptiles and amphibians in south-western Ecuador (Arteaga et al., 2016), and it lies close to the phylogeographical break of P. aequatorialis (Figure 1b). Nevertheless, individuals that belong to the northern lineage were sampled south of the river, indicating that the intraspecific divergence in P. aequatorialis is not caused by this geographic barrier. Such north–south phylogeographic breaks have been reported in other tropical rain forests because of differences in past and current climate and in geological history, including those from western Amazonia (Escobar et al., 2018; Jimenez-Vasquez et al., 2017; Reis et al., 2020; Roncal et al., 2015; Symula et al., 2003), eastern Brazil (Mousalli et al., 2009), central Africa (Bryja et al., 2014; Faye et al., 2016; Hardy et al., 2013; Helmstetter et al., 2020; Heuertz et al., 2014; Ley et al., 2017; Piñeiro et al., 2017) and Australia (Schneider et al., 1998). Analysing the phylogeographic patterns of more plant species will help to elucidate whether a phylogeographic break in south-western Ecuador is common for plants in the region.

The two evolutionary lineages detected for P. aequatorialis diverged 1.14 Ma (95% HPD: 1.01 – 2.48) during the Pleistocene (Figure 2), a period when glacial cycles impacted the distribution and genetic diversity of species (Hewitt, 2000, 2004). Based on our genetic and palaeoecological niche analyses (Figure 3c), P. aequatorialis potentially took refuge in two areas of historical climatic suitability in western Ecuador during the Pleistocene. The spatial distribution of the recovered genetic lineages overlaps with that of the two inferred potential refugia (Figure 3c), indicating that these areas of historical climatic suitability could have played a vicariant role in P. aequatorialis. Populations from northern and southern lineages apparently became isolated along the narrow strips of the Andean foothills of south-western Ecuador ~1 Ma during a glacial cycle and developed distinct genetic footprints. Gene flow between northern and southern populations could have occurred during the following interglacial cycles, hindering speciation between the two genetic lineages as suggested by their low level of genetic differentiation ($F_{ST} = 0.023$). The low admixture detected for the three northernmost individuals of the southern lineage (Figure 1c) partially supports the existence of gene flow between the two lineages. It is possible that stronger secondary contact has occurred within areas not sampled here. Therefore, detailed analyses of the populations located between the recovered distributions of the two genetic lineages will help to elucidate whether gene flow between the lineages has impeded...
speciation. Continuous isolation and secondary contact between the two lineages probably happened several times during the last 1 Ma. These contraction/expansion dynamics have promoted intraspecific divergence in *P. aequatorialis* during the Pleistocene. This has been observed in plants and animals from other tropical (Cabanne et al., 2016; Carnaval & Bates, 2007; Carnaval et al., 2009; Diniz-Filho et al., 2016; Faye et al., 2016; Hardy et al., 2013; Helmstetter et al., 2020; de Lima et al., 2014; Thomas et al., 2012; Vasconcellos et al., 2019) and temperate regions (Blanco-Pastor et al., 2019; Park & Donoghue, 2019).

The northern genetic lineage apparently took refuge along the humid lowlands and on the small mountain ranges of the Pacific coast, but not along the Andean foothills (Figure 3c). The inferred refugium for the northern lineage overlaps with the southern section of the suggested Chocó refugium (Gentry, 1982; Prance, 1982), which could have extended further south to the Pacific coast of western Ecuador. Our palaeoclimatic analyses suggest that the populations from the Andean foothills of north-western Ecuador migrated down-slope during glacial cycles (Figure 3), in agreement with the 'moist forest model' of montane and premontane rain forest evolution (Ramírez-Barahona & Eguiarte, 2013). Given that the Pacific coast remained moist at the time (Hooghiemstra & Van Der Hammen, 1998), the populations from the Andean foothills of north-western Ecuador were able to spread into the lowlands and potentially maintain connectivity with the populations from the Chocó refugium. Similarly, palynological data showed that plants from the eastern Ecuadorian Andean foothills descended ~1,000 m during glacial cycles (Brunschön & Behling, 2010; Cárdenas et al., 2011) when western Amazonia remained humid and covered by rain forest (Colinvaux et al., 2000). High palaeoclimatic suitability in coastal regions has also been detected in the Brazilian Atlantic region (Carnaval et al., 2009; Mousalli et al., 2009) and in central Africa (Faye et al., 2016; Helmstetter et al., 2020), indicating that coastal areas might have been important refugia for tropical rain forests.
We also inferred that the southern genetic lineage potentially took refuge in a narrow strip of Andean foothills in south-western Ecuador over 1,000 m.a.s.l. (Figure 3c; Table S1). These populations could have survived the Pleistocene isolated on the slopes of the Andes because they were unable to migrate down-slope into the lowlands of south-western Ecuador, which remained arid during recent glacial cycles (Heusser & Shackleton, 1994). Thus, the southern lineage evolved into a distinct genetic group following isolation as shown by our phylogenomic (Figure 1a; Figure S1) and population structure analyses (Figure 1b, 1c). In contrast with the northern lineage, these results support the ‘dry refugia model’ (Ramírez-Barahona & Eguiarte, 2013) in the premontane rain forests of south-western Ecuador. Thus, the Pleistocene climatic fluctuations could have affected contiguous rain forests differently depending on their latitudinal location. A similar response of the Andes may have occurred next to the dry inter-Andean valleys of Colombia where palynological data suggested that the Andean foothills constituted safe havens for rain forests (Bush et al., 2011; Hooghiemstra & Van Der Hammen, 2004). Rain forests also have remained stable throughout the Pleistocene in the Eastern Arc Mountains of Kenya and Tanzania where the lowlands were suggested to be dry during glacial cycles (Finch et al., 2009; Mumbi et al., 2008). Additional studies on the genetic composition and palaeoclimatic niche of plant and animal species are necessary to determine whether the Andean foothills of south-western Ecuador were indeed a refugium for rain forest species.

The reduced genetic diversity detected in the individuals from the Andean foothills of south-western Ecuador (Figure 4a), an area that potentially acted as a Pleistocene refugium, provides further evidence for the ‘dry refugia model’ acting on the Andean foothills of south-western Ecuador. These individuals may present lower genetic diversity compared to the individuals from the northern Andean foothills, which did not act as forest refugia, because of geographic isolation and potential population reduction during the Pleistocene (Ramírez-Barahona & Eguiarte, 2013). Nevertheless, it is also possible that this decrease in genetic diversity is due to differences in the number of individuals analysed and, therefore, these results should remain as exploratory. In addition, genetic diversity follows a geographic pattern by increasing jointly with decreasing latitude (Figure 4c). This suggests that P. aequatorialis could have colonized western Ecuador from the north because genetic diversity tends to decrease when species expand their ranges (Hewitt, 1999, 2000, 2004; Park & Donoghue, 2019; Petit et al., 2003; Zinck & Rajora, 2016). In contrast, the palaeoclimatic suitability index during the LGM did not influence genetic diversity in P. aequatorialis (Figure 4b), indicating that areas with more suitable palaeoclimatic conditions do not necessarily harbour higher current genetic diversity. A similar pattern was observed in the African rain forests for the palm Podococcus acaulis using unique haplotypes as a measure of genetic diversity (Faye et al., 2016). However, the sister species Podococcus barteri showed higher levels of unique genetic diversity in areas with high palaeoclimatic suitability during the LGM in the same rain forests (Faye et al., 2016). This suggests that the impact of climatic fluctuations on intraspecific diversity during the LGM varied between regions and even was different among closely related species within a same region. Thus, the current levels of genetic diversity in P. aequatorialis have been potentially shaped by the isolation of individuals in the Andean foothills during glacial cycles and by the species’ colonization history.

Phylogenomic relationships and geographical limits between P. aequatorialis and its sister species, P. tumacana, remain to be elucidated. Phytelephas tumacana occurs in south-western Colombia (Bernal & Galeano, 2010) at ~50 km from the three individuals of P. aequatorialis with which it formed a highly supported clade (Figure 1a; Figure S1). This suggests that P. tumacana could also occur in north-western Ecuador and we could have sampled it incorrectly as P. aequatorialis. There is also the possibility that we sampled a hybrid population between the two species, given the close proximity of their geographic ranges. One of the northernmost individuals of P. aequatorialis resolved with high support as a sister clade of the two intraspecific genetic lineages in the maximum likelihood tree (Figure S1), which adds uncertainty to the identity of P. aequatorialis populations in north-western Ecuador. In addition, the inability of our palaeoecological niche modelling to predict the current distribution of P. aequatorialis in north-western Ecuador suggests that different climatic conditions could be involved in its differentiation with P. tumacana. Detailed genetic, morphological and phenological studies must be performed to these populations to confirm their taxonomic identity.

5 | CONCLUSIONS

This is the first demonstration of the potential effects of Pleistocene climatic fluctuations on the distribution and genetic patterns of a plant species endemic to the rain forests of western Ecuador using nuclear genomic data. Our results support the role of Pleistocene climatic fluctuations and areas of historical climatic suitability in promoting intraspecific divergence in tropical species. Phytelephas aequatorialis shows a discrete north-south distribution of its intraspecific genetic lineages estimated to have originated ~1 Ma. This divergence was potentially caused by the isolation of populations into two rain forest refugia during glacial cycles, as suggested by the palaeoecological niche of P. aequatorialis during the LGM. The Chocó refugium could have extended further south to the Pacific coast of western Ecuador and could have included the populations of P. aequatorialis that descended from the northern Andean foothills. The Andean foothills of south-western Ecuador may be an important area for rain forest evolution because they could have remained climatically suitable through the Pleistocene. This promoted intraspecific divergence and decreased genetic diversity in the populations of P. aequatorialis from that region. Therefore, this study also provides phylogeographic and palaeoclimatic evidence for the ‘moist forest’ and ‘dry refugia’ models acting on the Andean foothills of western Ecuador (Ramírez-Barahona & Eguiarte, 2013), showing that contiguous
premontane rain forests could have been affected differently by Pleistocene climatic fluctuations. Our study shows that the rain forests of south-western Ecuador should be further studied and protected to preserve its unique intraspecific genetic diversity that has potentially originated during the Pleistocene.

ACKNOWLEDGEMENTS
We thank Ecuador’s Ministerio del Ambiente for providing the sampling permit to undertake this research (MAE-DNB-CM-2018-0082). We also thank Rodrigo Bernal for providing leaf material. This research was financially supported by the SENESCYT (PhD scholarship to S.E.), the Independent Research Fund Denmark (9040-00136B to H.B.), the Agence Nationale de la Recherche (ANR-15-CE02-0002-01 to T.L.P.C.) and the International Mixed Laboratory (LMI) BIO_INCA. We are also very grateful to the generous financial support provided by the International Palm Society (IPS) for funding a field trip to western Ecuador without which part of the sampling would not have been possible (field grant to S.E.). We dedicate this article to Dr. Calaway H. Dodson (1928-2020), botanist and orchidologist, who worked for many years on the flora of Ecuador and authored the five volumes of ‘Orchids of Ecuador’.

We also thank Rodrigo Bernal for providing leaf material. This research was supported by the Agence Nationale de la Recherche (ANR-15-CE02-0002-01 to T.L.P.C.) and the International Mixed Laboratory (LMI) BIO_INCA. We are also very grateful to the generous financial support provided by the International Palm Society (IPS) for funding a field trip to western Ecuador without which part of the sampling would not have been possible (field grant to S.E.). We dedicate this article to Dr. Calaway H. Dodson (1928-2020), botanist and orchidologist, who worked for many years on the flora of Ecuador and authored the five volumes of ‘Orchids of Ecuador’.

DATA AVAILABILITY STATEMENT
The FASTQ sequences (R1 and R2) for all individuals have been submitted to Genbank SRA under the BioProject PRJNA660607 https://www.ncbi.nlm.nih.gov/sra/PRJNA660607. Additional data relevant to this study can be found at Dryad https://doi.org/10.5061/dryad.98sf70mg2. Bioinformatic scripts used for this study can be found at https://github.com/ajhelmostetter/afrodyn and at https://github.com/ajhelmostetter/afrodyn.

ORCID
Sebastián Escobar https://orcid.org/0000-0002-4918-7967
Andrew J. Helmostetter https://orcid.org/0000-0003-3761-4981
Scott Jarvie https://orcid.org/0000-0002-0086-2351
Rommel Montúfar https://orcid.org/0000-0001-5309-4889
Henrik Balslev https://orcid.org/0000-0002-7101-7120
Thomas L.P. Couvreur https://orcid.org/0000-0002-8509-6587

REFERENCES
Abadi, S., Azoari, D., Pupko, T., & Mayrose, I. (2019). Model selection may not be a mandatory step for phylogeny reconstruction. Nature Communications, 10(1), https://doi.org/10.1038/s41467-019-08822-w
Acosta-Solis, M. (1948). Tagua or vegetable ivory—a forest product of Ecuador. Economic Botany, 2(1), 46–57. https://doi.org/10.1007/BF02907918
Aiello-Lammens, M. E., Boria, R. A., Radosavlievic, A., Vilela, B., & Anderson, R. P. (2015). spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography, 38(5), 541–545. https://doi.org/10.1111/ecog.01332
Andermann, T., Cano, Á., Zikza, A., Bacon, C., & Antonelli, A. (2018). SECAPR: A bioinformatics pipeline for the rapid and user-friendly processing of targeted enriched Illumina sequences, from raw reads to alignments. PeerJ, 2018(7), 1–15. https://doi.org/10.7717/peerj.5175
Arteaga, A., Pyron, R. A., Peñafliel, N., Romero-Barreto, P., Culebras, J., Bustamante, L., Yánez-Muñoz, M. H., & Guayasamin, J. M. (2016). Comparative phylogeography reveals cryptic diversity and repeated patterns of cladogenesis for amphibians and reptiles in Northwestern Ecuador. PLoS One, 11(4), e0151746. https://doi.org/10.1371/journal.pone.0151746
Barford, A. S. (1991). A monographic study of the subfamily Phylelephantoideae (Arecaceae). Opera Botanica, 105, 1–73.
Barford, A. S., Trénel, P., & Borchenius, F. (2010). Drivers of diversification in the vegetable ivory palms (Arecaceae: Ceroxylloideae, Phylelephloeae) – Vicariance or adaptive shifts in niche traits? In O. Seberg, G. Peterson, A. Barford, & J. Davis (Eds.), Diversity, Phylogeny, and Evolution in the Monocotyledons (pp. 225–243).
Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., Soberón, J., & Villalobos, F. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecological Modelling, 222(11), 1810–1819. https://doi.org/10.1016/j.ecolmodel.2011.02.011
Baumgartner, J. B., Esperón-Rodríguez, M., & Beaumont, L. J. (2018). Identifying in situ climate refugia for plant species. Ecography, 41(11), 1850–1863. https://doi.org/10.1111/ecog.03431
Behling, H., Hooghiemstra, H., & Negret, A. J. (1998). Holocene history of the Chocó rain forest from Laguna Piusbi, southern Pacific lowlands of Colombia. Quaternary Research, 50(3), 300–308. https://doi.org/10.1006/qres.1998.1998
Bernal, R., & Galeano, G. (2010). Palmas de Colombia: guía de campo. Panamerican Formas e Impressos S.A.
Blanco-Pastor, J. L., Manel, S., Barre, P., Roschanski, A. M., Willner, E., Dehmer, K. J., & Sampoux, J. P. (2019). Pleistocene climate changes, and not agricultural spread, accounts for range expansion and admixture in the dominant grassland species Lolium perenne L. Journal of Biogeography, 46(7), 1451–1465. https://doi.org/10.1111/jbi.13587
Borchenius, F., Borgoft Pedersen, H., & Balslev, H. (1998). Manual to the Palms of Ecuador. AUA Reports 37.
Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavrushkina, A., Heled, J., Jones, G., Kühnert, D., de Maio, N., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., ... Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Computational Biology, 15(4), 1–28. https://doi.org/10.1371/journal.pcbi.1006650
Brown, J. W., Walker, J. F., & Smith, S. A. (2017). Phyx: Phylogenetic tools for unix. Bioinformatics, 33(12), 1886–1888. https://doi.org/10.1093/bioinformatics/btx063
Brunschön, C., & Behling, H. (2010). Reconstruction and visualization of upper forest line and vegetation changes in the Andean depression region of southeastern Ecuador since the Last Glacial Maximum – A multi-site synthesis. Review of Palaeobotany and Palynology, 163(1–2), 139–152. https://doi.org/10.1016/j.revpalbo.2010.10.005
Bryja, J., Mikula, O., Patzenhauerová, H., Ogue, N. O., Šumbera, R., & Verheyen, E. (2014). The role of dispersal and vicariance in the Pleistocene history of an East African mountain rodent, Praomys delecorum. Journal of Biogeography, 41(1), 196–208. https://doi.org/10.1111/j.112195
Bush, M. B. (1994). Amazonian speciation: A necessarily complex model. Journal of Biogeography, 21(1), 5. https://doi.org/10.2307/2845600
Bush, M. B., Hanselman, J. A., & Hooghiemstra, H. (2011). Andean montane forests and climate change. In M. Bush, J. Flenley, & W. Gosling (Eds.), Evolution in the Dominant Grassland Species Lolium perenne L. Journal of Biogeography, 46(7), 1451–1465. https://doi.org/10.1111/jbi.13587
Cabanee, G. S., Calderón, L., Trujillo Arias, N., Flores, P., Pessoa, R., d’Horta, F. M., & Miyaki, C. Y. (2016). Effects of Pleistocene
climate changes on species ranges and evolutionary processes in the Neotropical Atlantic Forest. *Biological Journal of the Linnean Society*, 119(4), 856–872. https://doi.org/10.1111/bij.12844

Cárdenas, M. L., Gosling, W. D., Sherlock, S. C., Poole, I., Pennington, R. T., & Mothes, A. (2011). The response of vegetation on the Andean Flank in Western Amazonia to Pleistocene Climate Change. *Science*, 333(6051), 1055–1059. https://doi.org/10.1126/science.1207888

Carnaval, A. C., & Bates, J. M. (2007). Amphibian DNA shows marked genetic structure and tracks Pleistocene Climate change in northeastern Brazil. *Evolution*, 61(12), 2942–2957. https://doi.org/10.1111/j.1558-5646.2007.00241.x

Carnaval, A. C., Hickerson, M. J., Haddad, C. F. B., Rodrigues, M. T., & Moritz, C. (2009). Stability predicts genetic diversity in the Brazilian Atlantic Forest Hotspot. *Science*, 323(5915), 785–789. https://doi.org/10.1126/science.1166955

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, 17(4), 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334

Ceron, C., Palacios, W. A., Valencia, R., & Sierra, R. (1999). Propuesta preliminar de un sistema de clasificación de vegetación para el Ecuador. In R. Sierra (Ed.), *Las Formaciones Naturales de la Costa del Ecuador*. doi:https://doi.org/10.13140/2.1.4520.9287

Chamberlain, S., & Boettiger, C. (2017). R Python, and Ruby clients for a Neotropical savanna tree. *Ewédjé, E.-E., Gomez, C., Heuertz, M., Koffi, G. K., Lowe, A. J., Micheneau, C., Ndiade-Bourobou, D., Piñeiro, R., & Poncet, V. (2013). Molecular Biology and Evolution, 17(4), 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334*
Comparative phylogeography of African rain forest trees: A review of genetic signatures of vegetation history in the Guineo-Congolian region. *Comptes Rendus Geoscience*, 345(7–8), 284–296. https://doi.org/10.1016/j.crte.2013.05.001

Heine, K. (2000). Tropical South America during the Last Glacial Maximum: Evidence from glacial, periglacial and fluvial records. *Quaternary International*, 72(1), 7–21. https://doi.org/10.1016/S0149-0419(00)00017-3

Helmstetter, A. J., Amoussou, B. E. N., Bethune, K., Kandem, N. G., Glélé Kakai, R., Sonké, B., & Couvreur, T. L. P. (2020). Phylogenomic approaches reveal how climate shapes patterns of genetic diversity in an African rain forest tree species. *Molecular Ecology*, 29(18), 1–14. https://doi.org/10.1111/mec.15572

Heuertz, M., Duminil, J., Dauby, G., Savolainen, V., & Hardy, O. J. (2014). Comparative phylogeography in rainforest trees from Lower Guinea, Africa. *PLoS One*, 9(1), e84307. https://doi.org/10.1371/journ.al.pone.0084307

Heusser, L. E., & Shackleton, N. J. (1994). Tropical climatic variation on western Amazonia underlie the divergence of *Astrocaryum section Huicungo* (Arecaceae) and evaluation of potential anatomical adaptations. *Botanical Journal of the Linnean Society*, 185(3), 291–306. https://doi.org/10.1093/botlinnean/box060

Johnson, M. G., Gardner, E. M., Liu, Y., Medina, R., Goffinet, B., Shaw, A. J., Zerega, N. J. C., & Wickett, N. J. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences*, 4(7), 1600016. https://doi.org/10.3732/apps.1600016

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, 11(1), 94. https://doi.org/10.1186/1471-2156-11-94

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780. https://doi.org/10.1093/molbev/msm010

Leite, Y. L. R., Costa, L. P., Loss, A. C., Rocha, R. G., Batalha-Filho, H., Bastos, A. C., Quaresma, V. S., Fagundes, V., Paresque, R., Passamani, M., & Pardini, R. (2016). Neotropical forest expansion during the last glacial period challenges refuge hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 113(4), 1008–1013. https://doi.org/10.1073/pnas.1513062113

Ley, A. C., Heuertz, M., & Hardy, O. J. (2017). The evolutionary history of central African rain forest plants: Phylogeographical insights from sister species in the climber genus *Haumania* (Marantaceae). *Journal of Biogeography*, 44(2), 308–321. https://doi.org/10.1111/jbi.12902

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. *Bioinformatics*, 27(21), 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28(2), 298–299. https://doi.org/10.1093/bioinformatics/btr642

Liu, C., Newell, G., & White, M. (2016). On the selection of thresholds for predicting species occurrence with presence-only data. *Ecology and Evolution*, 6(1), 337–348. https://doi.org/10.1002/ece3.1878

Liu, C., White, M., & Newell, G. (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, 40(4), 778–789. https://doi.org/10.1111/jbi.12058

Maître, B. S., Boyle, B., Casler, N., Condit, R., Donoghue, J., Durán, S. M., & Enquist, B. J. (2018). The BIEN R package: A tool to access the botanical information and ecology network (BIEN) database. *Methods in Ecology and Evolution*, 9(2), 373–379. https://doi.org/10.1111/2041-210X.12861

Mialac, C., Scarcelli, N., Pouzadou, J., Barnaud, A., Billot, C., Faye, A., Kougbeadjo, A., Mailil, V., Martin, G., Sabot, F., Santoni, S., Vigouroux, Y., & Couvreur, T. L. P. (2014). Cost-effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies. *Molecular Ecology Resources*, 14(6), 1103–1113. https://doi.org/10.1111/1755-0998.12258

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBioJournal*, 17(1), 10. https://doi.org/10.14806/ej.17.1.200

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kerns, T., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for...
and basic results of CMIP5-20c3m experiments. Geoscientific Model Development, 4(4), 845–872. https://doi.org/10.5194/gmd-4-845-2011
Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics, 19(56), 153. https://doi.org/10.1186/s12859-018-2129-y
Zinck, J. W. R., & Rajora, O. P. (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), 1–17. https://doi.org/10.1186/s12862-016-0624-1
Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D., Antonelli, A., Farooq, H., Herdean, A., Ariza, M., Scharn, R., Svantesson, S., Wengström, N., Zizka, V., & Antonelli, A. (2019). CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. Methods in Ecology and Evolution, 10(5), 744–751. https://doi.org/10.1111/2041-210X.13152

BIOSKETCH
Sebastián Escobar conducted this research as part of his PhD program at the Department of Biology in Aarhus University. He is interested in unravelling the evolutionary history and phylogenetic relationships of tropical plants to understand the processes that have shaped current patterns of biodiversity.

Author contributions: R.M., H.B., and T.L.P.C conceived the study; S.E., S.J., R.M. and T.L.P.C. collected the data; S.E. performed laboratory work; S.E. analysed the data with assistance from A.J.H. and S.J.; S.E. led the writing with contribution from all authors.

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Escobar S, Helmstetter AJ, Jarvie S, Montúfar R, Balslev H, Couvreur TLP. Pleistocene climatic fluctuations promoted alternative evolutionary histories in Phytelephas aequatorialis, an endemic palm from western Ecuador. J Biogeogr. 2021;48:1023–1037. https://doi.org/10.1111/jbi.14055