Botanical Journal of the Linnean Society, 2017, **185**, 291–306. With 5 figures.

Dry season characteristics in western Amazonia underlie the divergence of *Astrocaryum* section *Huicungo* (Arecaceae) and evaluation of potential anatomical adaptations

VICTOR JIMENEZ-VASQUEZ[1,4*], BETTY MILLÁN[1,2], MIGUEL MACHAHUA[2], FRANCIS KAHN[2], RINA RAMIREZ[1,2], JEAN-CHRISTOPHE PINTAUD[3†] and JULISSA RONCAL[4]

[1]*Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Biológicas, Lima 1, Peru*
[2]*Universidad Nacional Mayor de San Marcos, Museo de Historia Natural, Apartado 14–0434, Lima 14, Peru*
[3]*UMR DIADE, Institut de Recherche pour le Développement, 34394 Montpellier, Cedex 5, France*
[4]*Department of Biology, Memorial University of Newfoundland, St John's, NL A1B 3X9, Canada*

Received 30 April 2017; revised 23 July 2017; accepted for publication 12 August 2017

Population and species divergence in South America are usually attributed to geographical barriers in the form of rivers, mountains or climate. In western Amazonia (<1000 m elevation) case studies addressing the ecological niche as a divergent selection agent are scarce. Using sequences from five plastid and six low-copy nuclear DNA regions, we reconstructed coalescent species phylogenetic trees for *Astrocaryum* section *Huicungo* (15 species, Arecaceae), which corroborated the presence of two lineages distributed north and south of 5°S in western Amazonia. Using elevation, three climatic and six soil variables we evaluated the ecological niche of each lineage. Different annual precipitation regimes were associated with each lineage. Notably, a lower precipitation seasonality and lower elevation were attributed for the northern clade and the opposite was found for the southern clade. We also explored the diagnostic and evolutionary importance of 35 anatomical and 31 morphological characters using a phylogenetic analysis and ancestral character reconstructions. None of the anatomical characters was diagnostic for either lineage. However, hypodermal cell wall width and the location of aerenchyma had different ancestral states for the two lineages, and their adaptive values to the dry season differences are discussed.

ADDITIONAL KEYWORDS: anatomy – coalescent species tree – dry season – ecological speciation – morphology – Neotropics – precipitation seasonality – species distribution modelling.

INTRODUCTION

Among the most studied abiotic factors that underlie phylogeographical breaks in South America are rivers, edaphic habitats, topography, climate and glaciation history (see Leite & Rogers, 2013; Turchetto-Zolet *et al.*, 2013 for reviews of phylogeographical studies in Amazonia and South America). In South America, the strongest phylogeographical break is attributed to the Andean uplift as a barrier (e.g. Cavers *et al.*, 2013; Bacon *et al.*, 2016). The Andean uplift has also been proposed to have facilitated diversification of plant lineages through climatic niche differentiation along an elevation–temperature gradient (e.g. Sanin *et al.*, 2016). Rivers may also act as dispersal barriers promoting population divergence (e.g. Fernandes, Wink & Aleixo, 2012; de Lima *et al.*, 2014; Fouquet *et al.*, 2015; Rocha *et al.*, 2015; but see Pomara, Ruokolainen & Young, 2014; Tuomisto *et al.*, 2016; Nazareno, Dick & Lohmann, 2017, for evidence of river permeability). Edaphic heterogeneity has been proposed as a driver of plant speciation in the Amazon basin (Fine *et al.*, 2005). For example, phylogenetic analyses in

*Corresponding author. E-mail: vr.jimenez.vs@gmail.com
†In memoriam.

Burseraceae tribe Protieae have demonstrated convergent evolution of soil associations, suggesting that adaptive evolution played an important role in the speciation of this group (Fine *et al.*, 2005). Species also segregate across environmental gradients or climatic discontinuities in tropical America. For example, seasonally dry forests have restricted seed dispersal of *Ficus insipida* Willd. between Meso- and South America (Honorio Coronado *et al.*, 2014). In the lizard *Anolis punctatus* a phylogeographical structure was observed for western, central Amazonia and the Atlantic coastal forest (Prates *et al.*, 2016).

Western Amazonia (WA) comprises the tropical rain forests from the eastern Andean foothills <500 m in elevation to the adjacent lowlands extending politically into the Amazonian regions of Colombia, Peru, Ecuador, north-western Bolivia and the states of Amazonas and Acre in western Brazil (figure 1 in Montufar & Pintaud, 2006). This region is of interest because of the large biodiversity it contains (ter Steege *et al.*, 2006) and because its geological history is much more recent and dynamic than that of eastern Amazonia (Rossetti, de Toledo & Goes, 2005). WA is composed of numerous sedimentary units deposited during the late Cenozoic as a consequence of the Andean uplift, which caused several episodes of subsidence that prompted retreat of a lake system at the end of the Miocene and gave rise to the development of a fluvial system with an eastward flow (Hoorn, 1993; Hoorn *et al.*, 1995; Vonhof, Wesselingh & Ganssen, 1998; Rossetti *et al.*, 2005). Shallow and short-lived marine incursions covered WA at least twice during the Miocene (Jaramillo *et al.*, 2017). Eastern Amazonia, by contrast, composed of old Palaeozoic and Precambrian rocks, remained geologically stable and had limited sediment deposition after the late Cretaceous (Aleixo & de Fatima Rossetti, 2007). In WA, case studies addressing abiotic factors underlying population or species divergence that may have acted as divergent selection agents are scarce.

Our case study, *Astrocaryum* G.Mey. section *Huicungo* F.Kahn (Arecaceae), is a clade of 15 species, of which 14 are present in WA, with most of them endemic to this region (Kahn, 2008; Roncal *et al.*, 2013). Three species reach Central Amazonia and the Guiana Shield (*A. ferrugineum* F.Kahn & B.Millán, *A. murumuru* Mart. and *A. ulei* Burret; Fig. 1). The palms in this group occupy the understorey to subcanopy (15 m) of rain forests in a variety of habitats ranging from primary to disturbed forests and from swamps to well-drained terra firme,

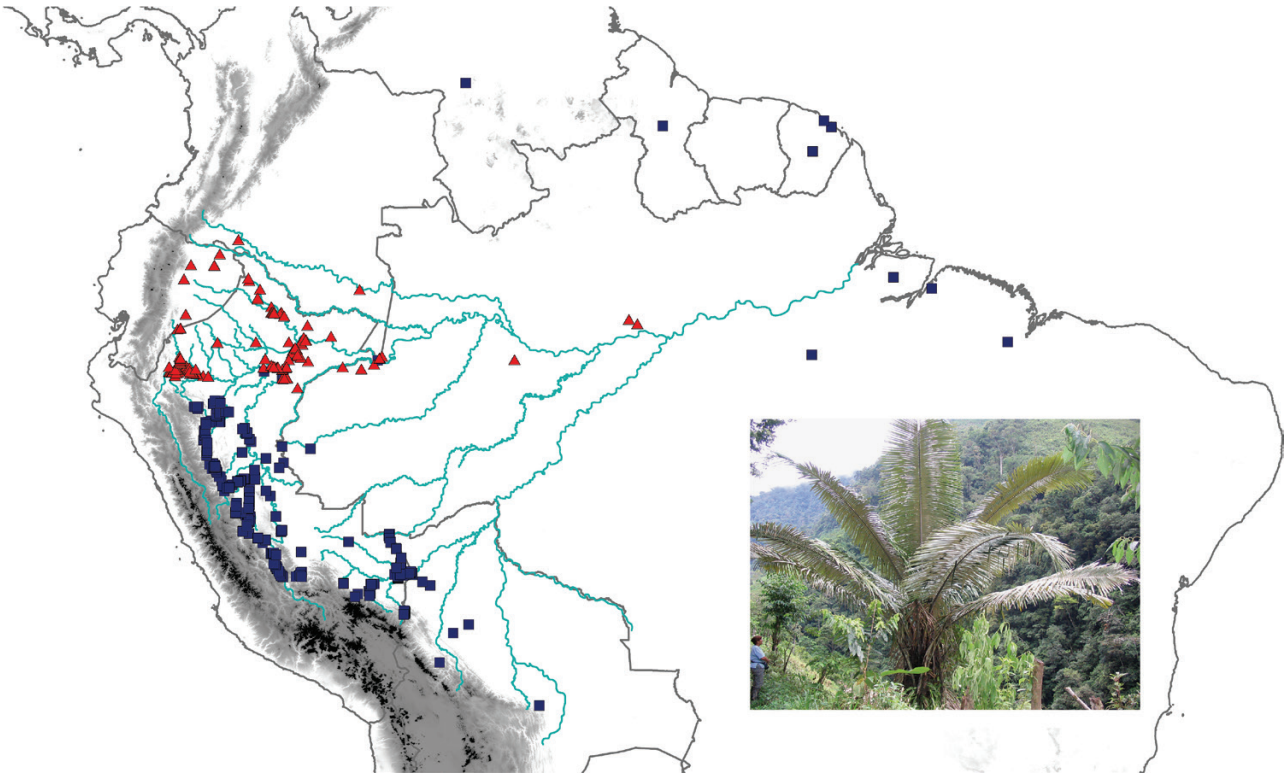

Figure 1. Map of northern South America showing occurrence data points for the central Andean foothills (blue squares) and northern Amazonian foreland basin (red triangles) clades used in the niche differentiation analysis. Inset photo of *Astrocaryum perangustatum* in Peru.

usually < 1000 m a.s.l. (rarely to 1200–1600 m). They tolerate deforestation and can be found in open areas. Detailed knowledge of the distribution of each species has been generated in Peru (Kahn *et al.*, 2011; Roncal *et al.*, 2015). The 'huicungo' palms, as rural inhabitants commonly call them, are not considered a commodity. However, local communities drink the liquid endosperm, the palm heart is edible, and oil extraction from the mesocarp and endocarp has been reported in Brazil (Kahn & Millán, 2013).

A synapomorphy of section *Huicungo* is that the rachillae or inflorescence branches have a glabrous proximal part, which is setose or pilose in the other three sections of subgenus *Monogynanthus* Burret (Kahn, 2008; Kahn & Millán, 2013). Species identification of the huicungo palms requires a combination of vegetative, pistillate flower and fruit morphological characters (Kahn & Millán, 2013). Anatomical characters of the lamina, main rib, petiole and sheath have also been proposed for taxonomic identification of *Astrocaryum* spp. (Millán & Kahn, 2010). Notably, characters of the stomata are distributed according to the infrageneric classification of *Astrocaryum* and the position of subsidiary terminal cells characterizes the different subsections in *Huicungo* (Millán & Kahn, 2010). However, morphological and anatomical characters have not been analysed in an evolutionary framework. The most recent study on evolutionary trends in leaf anatomy in palms did not include *Astrocaryum* (Horn *et al.*, 2009). Knowledge of the evolution of anatomical characters can provide insight on the selective regime (e.g. habitat type, climate) under which plant species evolved. For example, in the riparian *A. jauari* Mart., a well-developed aerenchyma transports gas from branches to roots under the hypoxic conditions of white- and black-water inundation forests, both leaf sides are heavily covered with wax to protect leaves from inundation and stomata are located only on the abaxial leaf surface (Schluter, Furch & Joly, 1993).

Cladogenesis of the huicungo palms was estimated at 6–8 Mya (Roncal *et al.*, 2013, 2015), post-dating the drainage of the Pebas aquatic system hypothesized to have dominated the WA landscape from 23 to 8 Mya (Wesselingh & Salo, 2006) and supporting its constraining role for colonization and *in situ* diversification. A time-calibrated tree and phylogeographical spatial diffusion analysis of the huicungo palms revealed a break in WA at *c.* 5°S between two main clades named the northern Amazonian foreland basin (NAFB) and central Andean foothills (CAF), with mean crown ages of 6.7 and 7.3 My, respectively (Roncal *et al.*, 2015). A biogeographical scenario was proposed of recurrent dispersal into WA from the Choco or the Guiana Shield followed by speciation in two areas of contrasting geological activity: tectonic uplift in the Fitzcarrald arch (FA) area in the CAF versus subsidence in the NAFB. It is hypothesized that uplift of the FA indicates the onset of rock uplift in the sub-Andean Shira mountains dated to after 9 ± 2 Mya and is related to subduction of the Nazca Ridge (Espurt *et al.*, 2010; Gautheron *et al.*, 2013). The FA played a dynamic role in modifying the Amazonian foreland basin because it rearranged drainage divides and catchment areas, erosional and sedimentological processes, generating a dynamic edaphic mosaic (Espurt *et al.*, 2010). These hydrological and sedimentological landscape modifications could have acted as a selection agent during the diversification of the huicungo palms, but the lack of detailed occurrence data prevented the evaluation of this hypothesis.

In the present study we test three hypotheses: (1) coalescence-based species tree reconstruction reveals the same two *Astrocaryum* section *Huicungo* lineages identified in a previous concatenation analysis; (2) ecological niche divergence might have played a role in the speciation of this group; and (3) leaf morphological and anatomical traits are useful to distinguish clades and evolved in conjunction with the ecological niche.

MATERIAL AND METHODS

Taxon sampling and DNA sequencing and alignment

We increased the number of individuals and DNA regions sequenced in Roncal *et al.* (2013, 2015) to improve the phylogenetic tree for the huicungo palms. We obtained silica-dried leaf material for 13 of the 15 species of *Astrocaryum* section *Huicungo* (*A. ciliatum* F.Kahn & B.Millán, and *A. cuatrecasanum* Dugand were unavailable). We used seven outgroup species representing sections *Astrocaryum* (*A. malybo* H.Karst, *A. vulgare* Mart.), *Euchambira* F.Kahn (*A. chambira* Burret), *Munbaca* Drude (*A. gynacanthum* Mart.) and *Monogynanthus* Burret [*A. sciophilum* (Miq.) Pulle] and two species of the sister genus *Hexopetion* Burret [*H. alatum* (H.F.Loomis) F.Kahn & Pintaud, *H. mexicanum* (Liebm. ex Mart.) Burret]. We followed the taxonomy of Kahn (2008). DNA was extracted from 25–45 µg of silica-gel-dried leaf tissue using the DNeasy plant mini kit (Qiagen, Valencia, CA, USA). We sequenced five plastid regions (the *psbM-trnD*, *rps16-trnQ* and *trnD-trnT* intergenic spacers and the *rps16* and *trnG* introns) for seven *Astrocaryum* individuals, a WRKY transcription factor (*WRKY7*) for 13 individuals, intron 4 of the phosphoribulokinase gene (*PRK*) for five individuals and the AGAMOUS paralogue 1 identified in the palm *Elaeis guineensis* Jacq. (*AG1*) for 39 individuals of *Astrocaryum* and *Hexopetion* (Supporting Information, Table S1). Primer pairs are specified in Table 1. Amplifications

Table 1. Primer pairs and annealing temperatures used to generate new DNA sequences for this study; alignment size in parentheses

	Locus (size)	Primer name	Annealing temperature (°C)	Reference
Nuclear	*WRKY7* (719 bp)	CnW07-F4 CnW07-R4	61	Meerow *et al.* (2009)
	PRK (1023 bp)	717F 969R	57	Lewis & Doyle (2002)
	AG1 (409 bp)	AG1-F1 AG1-R1	60	Ludeña et al. (2011)
Plastid	*rps16* intron (777 bp)	rps16-In-F rps16-In-R	58	Scarcelli *et al.* (2011)
	psbM-trnD (747 bp)	psbM-F trnD-R	53	Shaw *et al.* (2005)
	trnD-trnT (753 bp)	trnD-IGS-F trnT-IGS-R	62	Scarcelli *et al.* (2011)
	trnG intron (642 bp)	trnGin-F trnGin-R	61	Scarcelli *et al.* (2011)
	trnQ-rps16 (1086 bp)	trnQ-F rps16-R	62	Hahn (2002)

were performed on a 25-µL reaction mix containing 1 µL 1× Taq Buffer (Thermo Fisher Scientific, Waltham, MA, USA), 0.2 mM each dNTP, 0.2 µM each primer, 0.5 units PCR Taq DNA polymerase (Thermo Fisher Scientific), 1.25 mM $MgCl_2$ and *c.* 20 ng template DNA. PCR conditions consisted of 35 cycles with the following steps: denaturation at 94 °C for 1 min, annealing at various temperatures for 1 min (Table 1) and extension at 72 °C for 1 min. These cycles were preceded by an initial denaturation at 94 °C for 7 min and followed by a final extension at 72 °C for 10 min. PCR products were sent for Sanger sequencing to Macrogen Inc. (Seoul, Korea). In addition, the nuclear low-copy sequences of intron 23 of RNA polymerase II subunit 2 (*RPB2*), the region amplified by the conserved intron-scanning primer set #4 (*CISP4*) and phytochrome B (*PHYB*) were taken from Roncal *et al.* (2013). Specimens, herbarium voucher information and NCBI accession numbers are listed in Table S1 (Supporting Information). Each of the 11 DNA regions (five plastid and six nuclear) was independently aligned with MUSCLE (Edgar, 2004), manually adjusted in MEGA7 (Kumar, Stecher & Tamura, 2016) and subsequently concatenated in a supermatrix using SequenceMatrix 1.7.8 (Vaidya, Lohman & Meier, 2011). We removed a 108-bp region in *AG1* and a 36-bp microsatellite motif in *PHYB* because they were difficult to align. The total supermatrix consisted of 8131 bp for 79 individuals (57 from section *Huicungo* and 22 outgroups), with one to six individuals sampled per species and with 16% missing data. This matrix is available on TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S21294).

Molecular phylogenetic analyses

We performed a phylogenetic analysis on the supermatrix DNA alignment using maximum likelihood (ML) in RAxML v8.2.4 (Stamatakis, 2014). The best partition scheme was found using PartitionFinder (Lanfear *et al.*, 2012). An ML tree was reconstructed with 100 runs under the GTR+G+I model with the method 'lazy subtree rearrangement' from 100 different parsimony-starting trees. Branch support was assessed using the non-parametric Shimodaira–Hasegawa-like implementation of the approximate likelihood-ratio test (SH-aLRT; Guindon *et al.*, 2010). We used SH-aLRT instead of the bootstrap methodology because it is an order of magnitude faster, is based on whole alignment likelihoods rather than resampling random characters with replacement and is less sensitive to weak phylogenetic signal (Anisimova *et al.*, 2011; Pyron, 2011).

To address tree discordance among gene genealogies and to infer a species tree from multiple individuals per species, we used three methods that use a multispecies coalescent model. These three methods gave us insight on tree topology uncertainty and were useful to test how different topologies can affect the anatomical character state reconstruction analysis (see below). First, we used the NJst (Liu & Yu, 2011) and STAR (Liu *et al.*, 2009) algorithms as implemented in STRAW (Shaw *et al.*, 2013) with default parameters. These methodologies infer a species tree based on unrooted and rooted gene trees, respectively, and incorporate gene tree uncertainty using bootstrap analysis (Shaw *et al.*, 2013). For each of the six linkage units (all plastid regions were merged into one,

and *RPB2* was eliminated because it gave low-resolution trees), we generated 1000 bootstrap replicates with RAxML 8.2.4 under the GTR+G+I model. The 6000 bootstrap ML trees were compressed into a single zip file and uploaded into STRAW. We used NJst because it can handle multiple allele sequence data if the alleles sampled from the same species do not form a monophyletic group, which we demonstrated in the concatenation analysis (Liu & Yu, 2011). The STAR method was also performed because it is resistant to variable substitution rates along the branches in gene trees, and has been shown to outperform other coalescent methods such as STEAC, SC and GLASS (Liu *et al.*, 2009).

Second, because STAR and NJst may suffer from information loss and reduced efficiency, we performed a Bayesian analysis of the 11 DNA region dataset using the coalescent method of *BEAST (Heled & Drummond, 2010) as implemented in BEAST v1.8.2 (Drummond *et al.*, 2012) via the CIPRES Science Gateway (Miller, Pfeiffer & Schwartz, 2010). We used the partition schemes and substitution model as obtained in PartitionFinder. A relaxed molecular clock with an uncorrelated lognormal distribution under the Yule process as divergence tree prior, with random starting trees for each partition, was used. We ran two independent Markov chain Monte Carlo runs each consisting of 100 million generations with parameters sampled every 1000 generations and a 25% burn-in. We combined independent runs using LogCombiner v1.8.2 and evaluated convergence by considering estimated sample size (ESS) values > 200 for all model parameters in Tracer v1.8.2. We obtained a maximum clade credibility species tree with TreeAnnotator v1.8.2 (Drummond *et al.*, 2012).

Ecological niche divergence

We compiled 507 occurrence records of huicungo palms from Kahn *et al.* (2011), field observations in Bolivia, Peru, Ecuador, Colombia and Brazil, and Tropicos and Reflora databases (Fig. 1). To analyse the presence of niche divergence between clades we first conducted a principal component analysis (PCA) using ten environmental predictors: three climatic variables from WorldClim (Hijmans *et al.*, 2005); elevation from the NASA Shuttle Radar Topographic Mission 90m Digital Elevation Database v4.1 at 250-m spatial resolution (Jarvis *et al.*, 2008); and six soil variables from SoilGrids1km (Hengl *et al.*, 2014). These variables were chosen because they were not highly correlated (Pearson correlation coefficient < 0.60). Environmental data were handled and extracted in ArcGIS v10.4 (ESRI, Redlands, CA, USA). We conducted the PCA on the log-transformed continuous variables in R (R Core Team, 2015).

In addition, we used the maximum entropy method implemented in MaxEnt v3.3.3k (Phillips, Anderson & Schapire, 2006) to model the spatial distribution of each clade. We used 135 and 372 occurrence records for the NAFB and CAF clades, respectively, and the same set of ten variables employed in the PCA. We rescaled the elevation layer to 1 km^2 to meet the spatial resolution of the climatic and soil layers. The raw MaxEnt output was chosen because it does not rely on post-processing assumptions. We ran 15 replicates using 75% of the records as training data, the remaining 25% for testing the model and five feature classes (linear, quadratic, product, threshold and binary). We used default parameters for the maximum number of iterations (500), background points of 10 000, regularization multiplier of 1 and convergence threshold of 0.00001. The resultant prediction map represents the probability of occurrence ranging from 0 (low suitability) to 1 (high suitability). The importance of each environmental predictor in the model was evaluated through permutation importance metrics and a jackknife test. We used the receiver operation characteristic (ROC) approach by calculating the area under the ROC curve (AUC) to characterize model performance. Finally, we measured niche overlap between the two huicungo clades using Hellenger distance *I* (Warren, Glor & Turelli, 2008) and Schoener's *D* (Schoener, 1968) statistics as implemented in ENMTools (Warren, Glor & Turelli, 2010). Both measures range from 0 to 1 for completely discordant and identical distributions, respectively.

Morphological and anatomical characters

We recoded and rescored the list of 109 binary (presence/absence) leaf anatomical characters from the lamina, main rib, petiole and sheath published by Millán & Kahn (2010) as multistate characters, giving a new total of 35 characters (Supporting Information, Table S2). We recoded these binary characters because often the 'absence' state might be non-homologous, rendering incorrect suggestions of common ancestry and similarity in morphologically dissimilar taxa (Jenner, 2002). In addition, we newly coded and scored 31 morphological characters obtained from Kahn & Millán (2013), mostly describing the trunk, spines, inflorescence and female flower shape (Supporting Information, Table S2). The combined anatomical and morphological dataset (Supporting Information, Table S3) for 14 species of *Astrocaryum* section *Huicungo* and 12 outgroups (including *Hexopetion*) were used to reconstruct a phylogenetic tree using maximum parsimony in PAUP* 4.0b10a (Swofford, 2002). We used a stepwise addition heuristic search with 1000 replicates of tree-bisection-reconnection (TBR), keeping ten trees per replicate, and with accelerated character

transformation (ACCTRAN). Additionally, TBR was applied to all the trees retained in memory. All characters were considered unordered and equally weighted. Node support was assessed with 1000 non-parametric bootstrap replicates and a 50% majority rule consensus tree was constructed. The data matrix and consensus tree are available on TreeBASE.

To explore the diagnostic and evolutionary importance of the anatomical characters, we estimated ancestral states by maximum likelihood using the 'ace' command in the R package 'ape' (Paradis, Claude & Strimmer, 2004). We tested three different models for transition probabilities between the states of each discrete character: equal rates (ER), symmetrical (SYM) and all rates different (ARD). A likelihood test (G statistics) between these models asks whether the difference in log likelihoods is large enough to lie at the rightmost tail of a chi-square distribution (Paradis, 2012). In every case, the chi-square *P*-value was not significant (> 0.05), suggesting that all three models were not significantly different; we therefore chose the simplest ER model for every ancestral state reconstruction. To test how different topologies can affect the anatomical character state reconstruction analysis, we used the NJst and *BEAST coalescent trees.

RESULTS

Molecular phylogenetic analyses

PartitionFinder found four optimal partitions for the 11 DNA loci: *AG1*+*CISP4*+*RPB2*; *PRK*+*WRKY7*; *PHYB*; and plastid. GTR+G+I was the best nucleotide substitution model for every partition according to the corrected Akaike information criterion. The topology obtained from the supermatrix ML analysis recovered section *Huicungo* as monophyletic (SH-aLRT = 99) and confirms the presence of two well-supported clades (Supporting Information, Fig. S1): the NAFB and CAF with SH-aLRT values of 80 and 89, respectively. The interspecific relationships remained unresolved in each clade since only *A. murumuru* was monophyletic (SH-aLRT = 97). SH-aLRT support values ranged from 72 to 99 and rendered most species para- or polyphyletic. Every outgroup species, however, was monophyletic.

The three coalescent species trees recovered section *Huicungo* as monophyletic with strong support [*BEAST posterior probability (pp) = 0.97, NJst bootstrap support (BS) = 88.1 and STAR BS = 90]. The NAFB and CAF clades were recovered in all trees, except for the NAFB, which was paraphyletic in the STAR tree (Fig. 2). Branch support in section *Huicungo* was mostly weak in all three coalescent trees (pp < 0.86, BS < 65). The three topologies differed slightly, but species members of each clade were the same, except for *A. chonta* Mart. which appeared in the CAF clade in the NJst, STAR and concatenated ML trees and in the NAFB clade in the *BEAST tree (Fig. 2). This constitutes the main difference between the *BEAST tree and the rest. The sister relationship between *A. urostachys* Burret and *A. macrocalyx* Burret was recovered in the three coalescent trees with strong to moderate support (pp = 0.67, BS NJst = 95, BS STAR = 97). The ESS scores in Tracer from the *BEAST analysis were all > 200. The gene trees from the six different linkage units imported in STRAW had polytomies, and the *PHYB* marker provided the most resolution.

Ecological niche divergence

The first two principal components summarized 66.6% of the environmental variance. Individuals from the NAFB and CAF clades mostly discriminated along a diagonal line 130° from the first component (Fig. 3). Component 1 was largely related to elevation (load of 0.8), which was correlated with the mean temperature of the driest quarter (Bio9). Individuals in the NAFB clade grow at lower elevations (mean = 164 m asl, SD = 67.3) and higher Bio9 (mean = 25.8 °C, SD = 0.45) than CAF individuals (mean elevation = 457 m asl, SD = 291; Bio9 mean = 23.8 °C, SD = 1.6; Supporting Information, Fig. S2).

The validation of each MaxEnt model indicated that they were significantly better than the random distributions because the AUC values were well supported (AUC test/training for NAFB were 0.977/0.981, and for CAF were 0.950/0.956). Inspection of the models for each clade shows a clear niche differentiation between them (Fig. 4). Their geographical distributions are largely allopatric since areas of high niche suitability (blue in Fig. 4) showed no overlap. The models showed little predicted distribution overlap between the two clades corresponding to areas of low niche suitability (orange in Fig. 4). The distribution model of the NAFB clade is mainly restricted to the Marañon and upper Amazonas river basin in Peru, north into eastern Ecuador and, to a lesser extent, in Colombia (Amazonas) and north-western Brazil (Amazonas); by contrast, the CAF is mainly distributed in the eastern Andean foothills of Peru and Bolivia and, to a lesser extent, in western Brazil (Acre, Amazonas). The low niche suitability values in Brazil and the Guiana shield might be the result of the small number of occurrence records available for *A. murumuru*, *A. ferrugineum* and *A. ulei*. Climatic variables contributed the most to the MaxEnt models (not elevation or soil characteristics). The most important climatic variables for fitting the models were precipitation seasonality (Bio15) and precipitation of the warmest quarter (Bio18, Supporting Information, Table S4). Individuals in the NAFB clade grow in regions of lower Bio15 (mean = 16.02, SD = 4.18) than those of the CAF clade

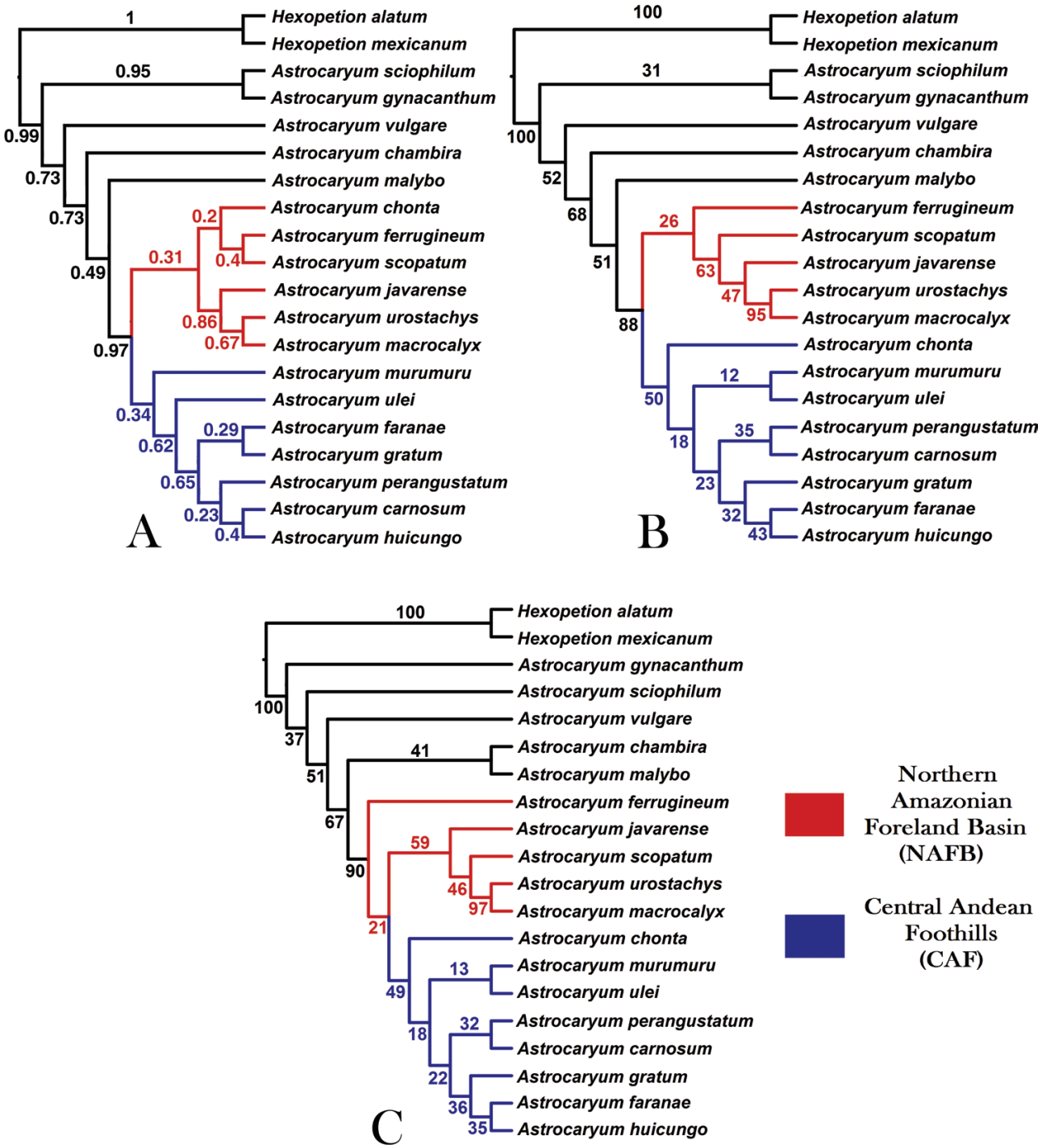

Figure 2. Species tree of *Astrocaryum* section *Huicungo* (Arecaceae) using three coalescent model methods based on five plastid and six low-copy nuclear DNA regions. Numbers along the branches are posterior probabilities for the *BEAST tree (A) and bootstrap support for NJst (B) and STAR (C) trees, respectively.

(mean = 41.67, SD = 9.87). Bio18, however, was not significantly different between clades. Since Bio15 was correlated with precipitation of the driest month (Bio14), we included this variable in the Supporting Information Fig. S2. A higher Bio14 (mean = 164.55, SD = 24.92) characterizes sites of the NAFB clade than those of the CAF clade (mean = 80.06, SD = 38.32). Niche overlap between clades was small to negligible, as revealed by Hellenger distance *I* of 0.2684 and Schoener's *D* of 0.0739.

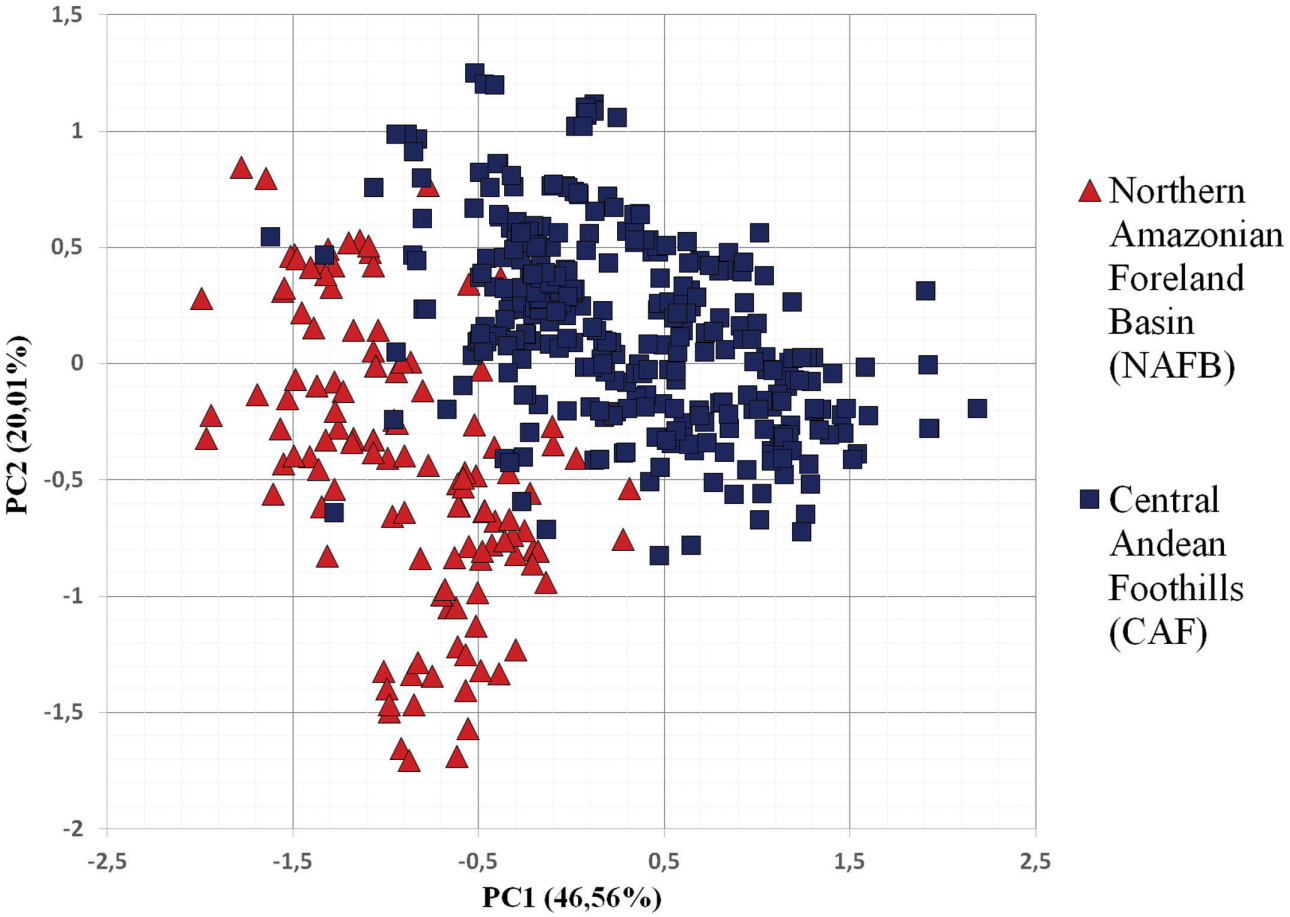

Figure 3. Principal component analysis of environmental characteristics (soil, climate and elevation) showing niche divergence of the central Andean foothills and northern Amazonian foreland basin lineages.

MORPHOLOGICAL AND ANATOMICAL CHARACTERS

The combined anatomical and morphological parsimony analysis resulted in a tree with low node support for the relationships among *Astrocaryum* (BS < 58, Supporting Information, Fig. S3). This analysis recovered *Astrocaryum* as monophyletic (BS = 98), but did not recover section *Huicungo*, or the NAFB or CAF clades. The sister relationship of *A. macrocalyx* and *A. urostachys* was recovered with weak support (BS = 45), and of *A. murumuru* and *A. ulei* with BS of 56. The analysis also recovered a group (*A. gratum* F.Kahn & B.Millán, *A. macrocalyx*, *A. perangustatum* F.Kahn & B.Millán, *A. urostachys*) compatible with subsection *Sachacungo* F.Kahn (Kahn, 2008) with weak support (BS = 34). The number of most-parsimonious trees was 23, tree length was 273, consistency index was 0.41 and retention index was 0.51, revealing high homoplasy.

No anatomical character was diagnostic for either clade, with most character reconstructions revealing the same most likely ancestral state for both clades at the time of divergence (e.g. hypodermal cell size, Fig. 5A). However, two drought-related anatomical characters showed different ancestral states for the two clades in the *BEAST tree. The ancestral condition for the hypodermal outer cell wall width on the abaxial side of leaves was thick for the CAF clade and thin for the NAFB clade (character A13, Fig. 5B). Aerenchyma was absent in the ancestor of a clade distributed in the eastern Andean foothills of Peru and Bolivia (CAF minus *A. murumuru*), and located in the centre of the main rib in the ancestor of the NAFB clade (character A28, Fig. 5C). When using the NJst tree, the ancestral states for the hypodermal outer cell wall width were equivocal for both clades. For the location of aerenchyma, the ancestor of the NAFB probably had no aerenchyma, whereas the reconstruction was equivocal for the ancestor of the CAF clade. We discuss the results of the *BEAST tree and disregard those of the NJst tree because of these multiple equivocal reconstructions for the nodes of interest. We showed that tree topology can affect the reconstruction of ancestral states, and thus our results need to be corroborated using a better-resolved species tree for section *Huicungo*.

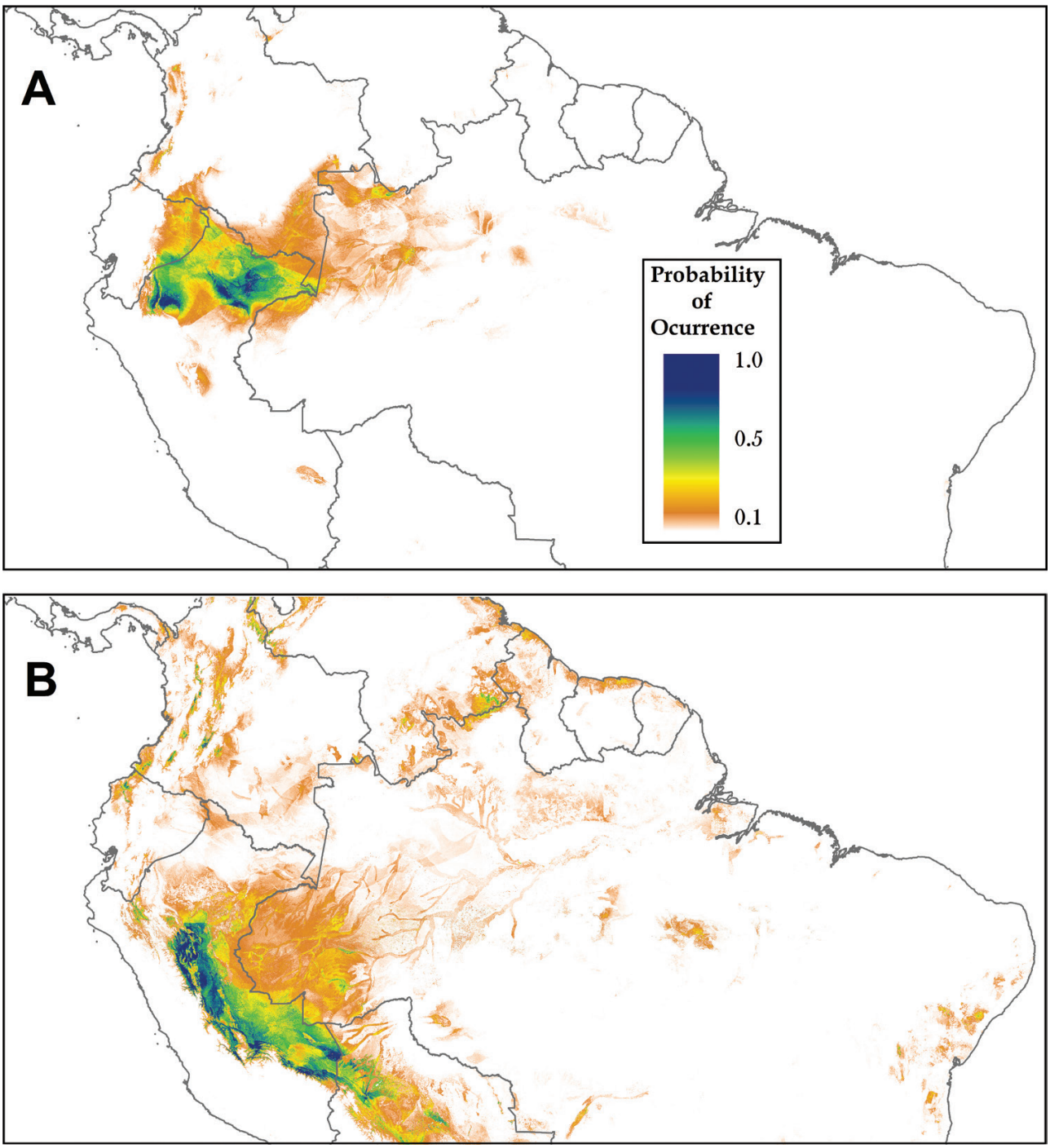

Figure 4. Species distribution modelling estimated using MaxEnt for (A) northern Amazonian foreland basin, and (B) central Andean foothills. Colours from orange via yellow and green to blue refer to MaxEnt values of occurrence probability (range 0–1, raw MaxEnt output).

DISCUSSION

MOLECULAR PHYLOGENETIC ANALYSES

Sequencing of 11 DNA loci allowed recognition of two clades, but resolution in each of them was still low. This was not surprising since no species in section *Huicungo* were monophyletic in the RAxML tree except for *A. murumuru*. Low resolution and support within clades are probably because of their recent origin (6–8 Mya, Roncal *et al.*, 2015), which might not

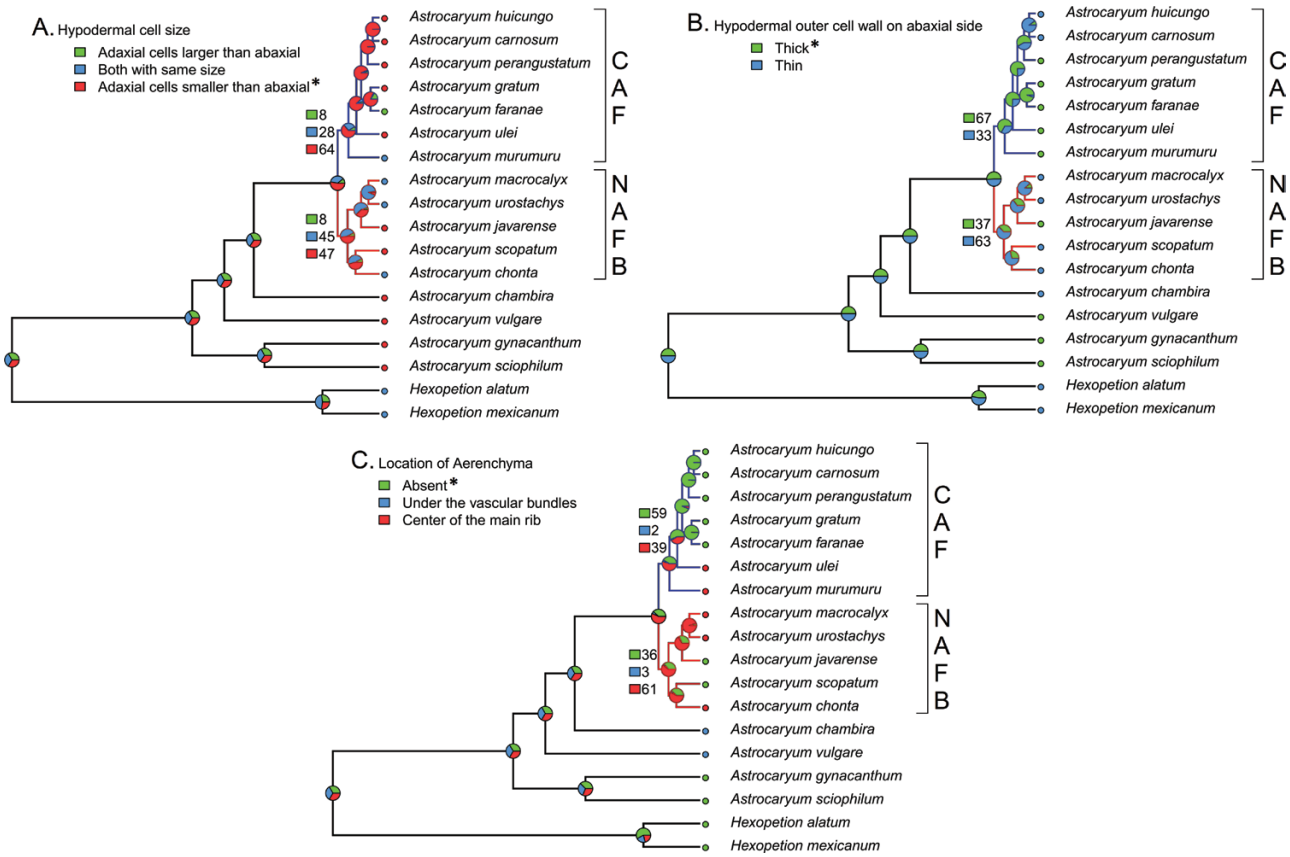

Figure 5. Ancestral character state reconstruction using maximum likelihood for three hypothesized drought-related anatomical characters plotted on the *BEAST coalescent tree. A, hypodermal cell size (character A10); B, hypodermal outer cell wall on abaxial side (character A13); C, aerenchyma (character A28). An asterisk marks states expected under dry conditions. Numbers next to boxes are the probabilities for each character state at the corresponding node.

have allowed the accumulation of numerous informative characters among species and the correct sorting of allelic variation. Alternatively, the evolution of the huicungo palms might not be entirely tree-like because of gene flow and hybridization, contributing to the degree of discordance we found among coalescent trees which only account for incomplete lineage sorting (ILS). The discordance among the three coalescent species trees in terms of some interspecific relationships can be explained by the sensitivity of these methods to the inherent lack of resolution in the gene trees (e.g. Huang *et al.*, 2010), which was evident from the polytomies recovered in all linkage unit trees (results not shown). The low bootstrap support of gene-tree-based methods (e.g. STAR, NJst) has been explained by gene tree estimation error, branch length information disregarded, and the addition of poorly supported gene trees (Liu *et al.*, 2015). Our results were concordant with claims that concatenation can result in more strongly supported phylogenetic trees than coalescent-based methods (e.g. Brumfield *et al.*, 2008; Townsend *et al.*, 2011), but the para- and polyphyly of species prevents us from relying solely on the concatenation approach.

Membership of clades does not match the classification of section *Huicungo* proposed by Kahn (2008). The clade of *A. murumuru*, *A. ulei* and *A. chonta* (subsection *Huicungo*) recovered by Roncal *et al.* (2015) did not appear in any phylogenetic analysis. Our results, therefore, do not support the taxonomic recognition of any subsection (*Sachacungo*, *Huicungo* or *Murumuru*) defined on the basis of pistillate flower morphological characters (Kahn, 2008) and anatomical stomata characteristics such as the position of subsidiary terminal cells (Millán & Kahn, 2010). Populations of *A. chonta* show a disjunct spatial distribution in northern and central Peru, which could explain its recovery in both clades in the coalescent trees. The sister relationship between *A. macrocalyx* and *A. urostachys* was consistently recovered in all molecular and morphological/anatomical phylogenetic analyses. These two species grow mostly in parapatry, adjacent to each other, with a small contact zone in some parts of the Napo and Tigre river valleys in Peru (Kahn & Millán, 2013).

Astrocaryum macrocalyx is found in terra firme forests from the north-eastern region of Peru in Loreto, northwards into Colombia, whereas *A. urostachys* is found in Restinga forests on periodically inundated alluvial soils, from the Tigre river valley in Peru, throughout Amazonian Ecuador and into adjacent Colombia (Kahn & Millán, 2013). An anatomical synapomorphy for these two sister species is the Y-shape of terminal cells in the stomata (Millán & Kahn, 2010).

As Pennington & Dick (2010) suggested, over the course of millions of years, long-distance seed dispersal may overwrite and obscure any genetic signature of specific events that may have acted as speciation drivers in Amazonia such as riverine barriers, tectonic arches or marine incursions. This concept of historical 'dispersal overwrite' may explain why some disjunct lineages appear closely related (Pennington & Dick, 2010). Our results did not show historical dispersal overwrite, at least during the early evolution of section *Huicungo*, because we found geographical structure at the deep node of the phylogenetic tree. Since most species of section *Huicungo* occur in allopatry (Kahn *et al.*, 2011; Roncal *et al.*, 2015), we expected to also find phylogenetic geographical structure at a finer scale (i.e. shallow nodes in the tree), but this was not found, and we attribute this to the lack of phylogenetic resolution rendered by the markers chosen and/or historical dispersal overwrite.

Ecological niche divergence

Ancestral populations can divide and radiate into novel ecological niches, which are then subject to divergent selection, leading to reproductive isolation and speciation (Nosil, 2012). This hypothesis of ecological speciation is consistent with phylogenetic trees showing their greatest ecological differentiation along the inner nodes (Losos, 2009), and we found this pattern in section *Huicungo*. The niche segregation we observed for the two *Huicungo* clades involved different precipitation regimes north and south of the *c.* 5°S break in WA, with a decreased precipitation and increased length of the dry season south of the break. We acknowledge that our niche analysis considered only current environmental data and not the conditions under which the evolution of section *Huicungo* occurred (6–8 Mya). However, recent modelling of the effects of lower elevation during Andean upheaval suggested that total precipitation remained similar for most of the Amazon basin during the Neogene, not altering rainfall patterns enough to depart from a wet and warm tropical climate (Vonhof & Kaandorp, 2010). Since the Andes acted as a moisture barrier for atmospheric circulation, Amazonian climate probably experienced enhanced seasonality during the Neogene (Sepulchre, Sloan & Fluteau, 2010) and therefore the current precipitation regime we suggest as having promoted niche differentiation was perhaps already present during the early divergence of section *Huicungo*. The interpretation that the niche differentiation found between clades acted as a divergent selection agent for ecological speciation awaits a formal test on the presence and strength of selection and gene flow.

Using forest inventory data, ter Steege *et al.* (2006) found two dominant geographical gradients in genus-level plant community composition across northern South America. The first community gradient was correlated with a gradient in soil fertility, whereas the second one paralleled a gradient in dry season length stretching from Colombia (least dry) to south-eastern Amazonia (figure 1b in ter Steege *et al.*, 2006). This second community gradient matches the spatial distribution of the two *Huicungo* clades quite well. Individuals in the CAF grow under an annual precipitation seasonality (variance) of almost three times the seasonality occurring in the NAFB. Likewise, precipitation of the driest month (correlated with seasonality) was half for individuals in the CAF, compared with NAFB individuals (Supporting Information, Fig. S2). Although studies like that of ter Steege *et al.* (2006) are useful to detect the relative performance of species with respect to the environment, they are insufficient to demonstrate that climate in general and dry season length specifically can drive speciation. A phylogenetic framework like ours needs to be conducted for a wider spectrum of taxa in WA.

Morphological and anatomical characters

Horn *et al.* (2009) argued that a functional significance could be attributable to many lamina anatomical characters, compensating for the morphological limitations in palms, and that such anatomical variation appears to be largely decoupled from lamina morphology. They further argued that evolution of palm leaf anatomy is homoplasious, but not at random since they observed recurrent evolution of character states. In our case study, we found multiple shifts among character states, with no anatomical character useful to distinguish the two *Huicungo* clades unequivocally. The parsimony analysis based on anatomical and morphological characters resulted in a poorly resolved tree with almost no similarity to the molecular trees, suggesting that the characters coded might be plastic or lack phylogenetic utility. Despite the fact that we tested some anatomical characters previously diagnosed as leaf adaptations that prevent water loss, only two characters showed ancestral states that were consistent with expectations if these were adaptations to the dry precipitation regimes in WA. The first was the thickness of the hypodermal cell wall on the leaf abaxial surface. Thickening of outer walls in the epidermis

and hypodermis has been associated with the osmotic adjustment that enables plants to survive under water stress (Bussotti *et al.*, 1998). As expected for individuals exposed to a pronounced dry season, the ancestor of the dry CAF clade most probably had thick hypodermal cell walls, whereas the ancestor of the wet NAFB clade probably had thin hypodermal cell walls.

The second character was the location of aerenchyma, which was reconstructed as absent for the ancestor of the dry CAF minus *A. mururmuru* clade, and at the centre of the main rib for the ancestor of the wet NAFB clade. A reduction of mesophyll intercellular spaces (aerenchyma) provides internal blocking of water movement reducing transpiration (Bosabalidis & Kofidis, 2002). A third character, hypodermal cell size, could not be reconstructed unequivocally for the ancestor of the wet NAFB clade, for which we expected to find an ancestor with equal cell size in the abaxial and adaxial leaf surfaces. However, the dry CAF ancestor did show the expected state of smaller hypodermal cells on the adaxial surface, which has been shown to prevent cell collapse under water stress conditions (Oertli, Lips & Agami, 1990).

Among the anatomical drought tolerance characters that did not differ for the two lineage ancestors were cell wall thickness of the epidermis, papillate or hairy indumentum and bundle sheath extension. Besides osmotic adjustment, cell wall thickening and lignification in the epidermis can be due to autophagic processes with exocytosis of degraded material in response to drought (Vollenweider *et al.*, 2016). A higher density of non-glandular scales or hairs on the leaf surface is known to block externally water vapour movement reducing transpiration (Bosabalidis & Kofidis, 2002). Bundle sheath extension refers to a group of cells (usually fibres) that bridge veins to the surface layers of leaves. They occur in many angiosperms and are characteristic of heterobaric leaves, which are known to confer adaptive benefits such as structural support under water stress (Kenzo *et al.*, 2007) and accelerated stomatal response to humidity (Buckley, Sack & Gilbert, 2011). We acknowledge that incomplete sampling (13 out of 15 species) and different topologies could affect the character state reconstruction results presented here. Thus, the drought-related potential anatomical adaptations we propose in this study need to be confirmed with a better resolved, taxonomically complete species tree.

Conclusions and future research prospects

This study is one of a few to address the climatic niche as a potential divergent selection agent in lowland WA plants. The distribution of precipitation in WA differentiates the ecological niche of the NAFB and CAF *Huicungo* clades. The anatomical characters evaluated were not diagnostic of these two clades. However, hypodermal cell wall thickness and location of aerenchyma might have played a role as adaptations to the dry season length at the initial split of section *Huicungo*.

The use of targeted sequence enrichment and restriction-site-associated DNA sequencing methods may increase the phylogenetic resolution. Hypothesis testing on ecological speciation at a finer geographical scale than that presented here and confirmation of the potential anatomical adaptations could be possible with this improved phylogenetic tree. The multispecies coalescent methods used in this study imply only ILS as the source of gene tree incongruence. It is well known, however, that ILS by itself cannot explain all gene tree differences, and thus the *Huicungo* DNA dataset awaits to be evaluated with other species tree methods that incorporate hybridization, gene flow and reticulation.

ACKNOWLEDGEMENTS

This work was funded by an NSERC-Discovery grant (RGPIN-2014–03976) to J.R., the Emerging Leaders in the Americas Scholarship offered by the Canadian Ministry of Foreign Affairs to V.J., and the Vicerectorate of Research of UNMSM (#141001031 and #151001021) to R.R. We are grateful to Yolanda Wiersma for discussions on niche modelling. We are also grateful to Christine Bacon, an anonymous reviewer and the editor for improving the quality of this work. This study constitutes part of the MSc thesis in Molecular Biology at UNMSM of V.J. and is dedicated to the memory of Jean-Christophe Pintaud and Gloria Galeano, two excellent palm researchers, mentors and friends.

REFERENCES

Aleixo A, de Fatima Rossetti D. 2007. Avian gene trees, landscape evolution, and geology: towards a modern synthesis of Amazonian historical biogeography? *Journal of Ornithology* **148:** S443–S453.

Anisimova M, Gil M, Dufayard JF, Dessimoz C, Gascuel O. 2011. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Systematic Biology* **60:** 685–699.

Bacon CD, Velasquez-Puentes F, Florez-Rodriguez A, Balslev H, Galeano G, Bernal R, Antonelli A. 2016. Phylogenetics of Iriarteeae (Arecaceae), cross-Andean disjunctions and convergence of clustered infructescence morphology in *Wettinia*. *Botanical Journal of the Linnean Society* **182:** 272–286.

Bosabalidis A, Kofidis G. 2002. Comparative effects of drought stress on leaf anatomy of two olive cultivars. *Plant Science* **163:** 375–379.

Brumfield RT, Liu L, Lum DE, Edwards SV. 2008. Comparison of species tree methods for reconstructing the phylogeny of bearded manakins (Aves: Pipridae, *Manacus*) from multilocus sequence data. *Systematic Biology* **57:** 719–731.

Buckley TN, Sack L, Gilbert ME. 2011. The role of bundle sheath extensions and life form in stomatal responses to leaf water status. *Plant Physiology* **156:** 962–973.

Bussotti F, Gravano E, Grossoni P, Tani C. 1998. Occurrence of tannins in leaves of beech trees (*Fagus sylvatica*) along an ecological gradient, detected by histochemical and ultrastructural analyses. *New Phytologist* **138:** 469–479.

Cavers S, Telford A, Arenal Cruz F, Perez Castaneda AJ, Valencia R, Navarro C, Buonamici A, Lowe AJ, Vendramin GG. 2013. Cryptic species and phylogeographical structure in the tree *Cedrela odorata* L. throughout the Neotropics. *Journal of Biogeography* **40:** 732–746.

de Lima NE, Lima-Ribeiro MS, Tinoco CF, Terribile LC, Collevatti RG. 2014. Phylogeography and ecological niche modelling, coupled with the fossil pollen record, unravel the demographic history of a Neotropical swamp palm through the Quaternary. *Journal of Biogeography* **41:** 673–686.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29:** 1969–1973.

Edgar R. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5:** 1–19.

Espurt N, Baby P, Brusset S, Roddaz M, Hermoza W, Barbarand J. 2010. The Nazca Ridge and uplift of the Fizcarrald Arch: implications for regional geology in northern South America. In: Hoorn C, Wesselingh F eds. *Amazonia: landscape and species evolution. A look into the past*. Oxford: Wiley-Blackwell, 89–100.

Fernandes AM, Wink M, Aleixo A. 2012. Phylogeography of the chestnut-tailed antbird (*Myrmeciza hemimelaena*) clarifies the role of rivers in Amazonian biogeography. *Journal of Biogeography* **39:** 1524–1535.

Fine PV, Daly DC, Villa Muñoz G, Mesones I, Cameron KM. 2005. The contribution of edaphic heterogeneity to the evolution and diversity of Burseraceae trees in the western Amazon. *Evolution* **59:** 1464–1478.

Fouquet A, Courtois EA, Baudain D, Lima JD, Souza SM, Noonan BP, Rodrigues MT. 2015. The trans-riverine genetic structure of 28 Amazonian frog species is dependent on life history. *Journal of Tropical Ecology* **31:** 361–373.

Gautheron C, Espurt N, Barbarand J, Roddaz M, Baby P, Brusset S, Tassan-Got L, Douville E. 2013. Direct dating of thick- and thin-skin thrusts in the Peruvian Subandean zone through apatite (U–Th)/He and fission track thermochronometry. *Basin Research* **25:** 419–435.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59:** 307–321.

Hahn WJ. 2002. A phylogenetic analysis of the Arecoid Line of palms based on plastid DNA sequence data. *Molecular Phylogenetics and Evolution* **23:** 189–204.

Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27:** 570–580.

Hengl T, de Jesus JM, MacMillan RA, Batjes NH, Heuvelink GB, Ribeiro E, Samuel-Rosa A, Kempen B, Leenaars JG, Walsh MG, Gonzalez MR. 2014. SoilGrids1km–global soil information based on automated mapping. *PloS One* **9:** e105992.

Hijmans R, Cameron S, Parra J, Jones P, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25:** 1965–1978.

Honorio Coronado EN, Dexter KG, Poelchau MF, Hollingsworth PM, Phillips OL, Pennington RT, Carine M. 2014. *Ficus insipida* subsp. *insipida* (Moraceae) reveals the role of ecology in the phylogeography of widespread Neotropical rain forest tree species. *Journal of Biogeography* **41:** 1697–1709.

Hoorn C. 1993. Marine incursions and the influence of Andean tectonics on the Miocene depositional history of northwestern Amazonia – results of a palynostratigraphic study. *Palaeogeography Palaeoclimatology Palaeoecology* **105:** 267–309.

Hoorn C, Guerrero J, Sarmiento G, Lorente M. 1995. Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. *Geology* **23:** 237–240.

Horn JW, Fisher JB, Tomlinson PB, Lewis CE, Laubengayer K. 2009. Evolution of lamina anatomy in the palm family (Arecaceae). *American Journal of Botany* **96:** 1462–1486.

Huang H, He Q, Kubatko LS, Knowles LL. 2010. Sources of error inherent in species-tree estimation: impact of mutational and coalescent effects on accuracy and implications for choosing among different methods. *Systematic Biology* **59:** 573–583.

Jaramillo C, Romero I, D'Apolito C, Bayona G, Duarte E, Louwye S, Escobar J, Luque J, Carrillo-Briceño JD, Zapata V, Mora A, Schouten S, Zavada M, Harrington G, Ortiz J, Wesselingh FP. 2017. Miocene flooding events of western Amazonia. *Science Advances* **3:** e1601693.

Jarvis A, Reuter H, Nelson A, Guevara E. 2008. *Hole-filled SRTM for the globe Version 4*, available from the CGIAR-CSI SRTM 90m Database. Available at: http://srtm.csi.cgiar.org (accessed June 2017).

Jenner R. 2002. Boolean logic and character state identity: pitfalls of character coding in metazoan cladistics. *Contributions to Zoology* **71:** 67–91.

Kahn F. 2008. The genus *Astrocaryum* (Arecaceae). *Revista Peruana de Biologia* **15:** 31–48.

Kahn F, Millán B. 2013. *Las palmeras huicungo - the huicungo palms*. Lima: Universidad Nacional Mayor de San Marcos - Institut de Recherche pour le Développement.

Kahn F, Millán B, Pintaud J, Machahua M. 2011. Detailed assessment of the distribution of *Astrocaryum* sect. *Huicungo* (Arecaceae) in Peru. *Revista Peruana de Biologia* **18:** 279–282.

Kenzo T, Ichie T, Watanabe Y, Hiromi T. 2007. Ecological distribution of homobaric and heterobaric leaves in tree species of Malaysian lowland tropical rainforest. *American Journal of Botany* **94:** 764–775.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33:** 1870–1874.

Lanfear R, Calcott B, Ho SY, Guindon S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* **29:** 1695–1701.

Leite RN, Rogers DS. 2013. Revisiting Amazonian phylogeography: insights into diversification hypotheses and novel perspectives. *Organisms Diversity & Evolution* **13:** 639–664.

Lewis C, Doyle J. 2002. A phylogenetic analysis of tribe Areceae (Arecaceae) using two low-copy nuclear genes. *Plant Systematics and Evolution* **236:** 1–17.

Liu L, Xi Z, Wu S, Davis CC, Edwards SV. 2015. Estimating phylogenetic trees from genome-scale data. *Annals of the New York Academy of Sciences* **1360:** 36–53.

Liu L, Yu L. 2011. Estimating species trees from unrooted gene trees. *Systematic Biology* **60:** 661–667.

Liu L, Yu L, Pearl DK, Edwards SV. 2009. Estimating species phylogenies using coalescence times among sequences. *Systematic Biology* **58:** 468–477.

Losos J. 2009. *Lizards in an evolutionary tree: ecology and adaptive radiation of anoles.* Berkeley: University of California Press.

Ludeña B, Chabrillange N, Aberlenc-Bertossi F, Adam H, Tregear JW, Pintaud JC. 2011. Phylogenetic utility of the nuclear genes AGAMOUS 1 and PHYTOCHROME B in palms (Arecaceae): an example within Bactridinae. *Annals of Botany* **108:** 1433–1444.

Meerow AW, Noblick L, Borrone JW, Couvreur TL, Mauro-Herrera M, Hahn WJ, Kuhn DN, Nakamura K, Oleas NH, Schnell RJ. 2009. Phylogenetic analysis of seven WRKY genes across the palm subtribe Attaleinae (Arecaceae) [corrected] identifies *Syagrus* as sister group of the coconut. *PloS One* **4:** e7353.

Millán B, Kahn F. 2010. Characterization of leaf anatomy in species of *Astrocaryum* and *Hexopetion* (Arecaceae). *Revista Peruana de Biologia* **17:** 81–94.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE)*. New Orleans: Institute for Electrical and Electronic Engineers (IEEE), 1–8.

Montufar R, Pintaud J. 2006. Variation in species composition, abundance and microhabitat preferences among western Amazonian terra firme palm communities. *Botanical Journal of the Linnean Society* **151:** 127–140.

Nazareno AG, Dick CW, Lohmann LG. 2017. Wide but not impermeable: testing the riverine barrier hypothesis for an Amazonian plant species. *Molecular Ecology* **26:** 3636–3648.

Nosil P. 2012. *Ecological speciation*. Oxford: Oxford University Press.

Oertli JJ, Lips SH, Agami M. 1990. The strength of sclerophyllous cells to resist collapse due to negative turgor pressure. *Acta Oecologica* **11:** 281–289.

Paradis E. 2012. *Analysis of phylogenetics and evolution with R, 2nd edn*. New York: Springer.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20:** 289–290.

Pennington RT, Dick CW. 2010. Diversification of the Amazonian flora and its relation to key geological and environmental events: a molecular perspective. In: Hoorn C, Wesselingh FP eds. *Amazonia, landscape and species evolution: a look into the past.* Oxford: Wiley-Blackwell, 373–385.

Phillips S, Anderson R, Schapire R. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190:** 231–259.

Pomara LY, Ruokolainen K, Young KR. 2014. Avian species composition across the Amazon River: the roles of dispersal limitation and environmental heterogeneity. *Journal of Biogeography* **41:** 784–796.

Prates I, Rivera D, Rodrigues MT, Carnaval AC. 2016. A mid-Pleistocene rainforest corridor enabled synchronous invasions of the Atlantic Forest by Amazonian anole lizards. *Molecular Ecology* **25:** 5174–5186.

Pyron RA. 2011. Divergence time estimation using fossils as terminal taxa and the origins of Lissamphibia. *Systematic Biology* **60:** 466–481.

R Core Team. 2015. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Rocha RG, Ferreira E, Loss AC, Heller R, Fonseca C, Costa LP. 2015. The Araguaia River as an important biogeographical divide for didelphid marsupials in central Brazil. *Journal of Heredity* **106:** 593–607.

Roncal J, Couderc M, Baby P, Kahn F, Millán B, Meerow AW, Pintaud J. 2015. Palm diversification in two geologically contrasting regions of western Amazonia. *Journal of Biogeography* **42:** 1503–1513.

Roncal J, Kahn F, Millán B, Couvreur TLP, Pintaud J. 2013. Cenozoic colonization and diversification patterns of tropical American palms: evidence from *Astrocaryum* (Arecaceae). *Botanical Journal of the Linnean Society* **171:** 120–139.

Rossetti D, de Toledo P, Goes A. 2005. New geological framework for western Amazonia (Brazil) and implications for biogeography and evolution. *Quaternary Research* **63:** 78–89.

Sanin MJ, Kissling WD, Bacon CD, Borchsenius F, Galeano G, Svenning J, Olivera J, Ramirez R, Trenel P, Pintaud J. 2016. The Neogene rise of the tropical Andes facilitated diversification of wax palms (*Ceroxylon*: Arecaceae) through geographical colonization and climatic niche separation. *Botanical Journal of the Linnean Society* **182:** 303–317.

Scarcelli N, Barnaud A, Eiserhardt W, Treier UA, Seveno M, d'Anfray A, Vigouroux Y, Pintaud JC. 2011. A set of 100 chloroplast DNA primer pairs to study population genetics and phylogeny in monocotyledons. *PloS One* **6:** e19954.

Schluter U, Furch B, Joly C. 1993. Physiological and anatomical adaptations by young *Astrocaryum jauari* Mart. (Arecaceae) in periodically inundated biotopes of central Amazonia. *Biotropica* **25:** 384–396.

Schoener T. 1968. *Anolis* lizards of Bimini: resource partitioning in a complex fauna. *Ecology* **49:** 704–726.

Sepulchre P, Sloan LC, Fluteau F. 2010. Modelling the response of Amazonian climate to the uplift of the Andean mountain range. In: Hoorn C, Wesselingh FP eds. *Amazonia, landscape and species evolution: a look into the past.* Oxford: Wiley-Blackwell, 211–222.

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL. 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany* **92:** 142–166.

Shaw TI, Ruan Z, Glenn TC, Liu L. 2013. STRAW: Species TRee Analysis Web server. *Nucleic Acids Research* **41:** W238–W241.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312–1313.

Swofford DL. 2002. *PAUP*4.0. Phylogenetic analysis using parsimony (*and other methods)*. Sunderland: Sinauer Associates.

ter Steege H, Pitman NC, Phillips OL, Chave J, Sabatier D, Duque A, Molino JF, Prévost MF, Spichiger R, Castellanos H, von Hildebrand P, Vásquez R. 2006. Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature* **443:** 444–447.

Townsend TM, Mulcahy DG, Noonan BP, Sites JW Jr, Kuczynski CA, Wiens JJ, Reeder TW. 2011. Phylogeny of iguanian lizards inferred from 29 nuclear loci, and a comparison of concatenated and species-tree approaches for an ancient, rapid radiation. *Molecular Phylogenetics and Evolution* **61:** 363–380.

Tuomisto H, Moulatlet GM, Balslev H, Emilio T, Figueiredo FOG, Pedersen D, Ruokolainen K. 2016. A compositional turnover zone of biogeographical magnitude within lowland Amazonia. *Journal of Biogeography* **43:** 2400–2411.

Turchetto-Zolet AC, Pinheiro F, Salgueiro F, Palma-Silva C. 2013. Phylogeographical patterns shed light on evolutionary process in South America. *Molecular Ecology* **22:** 1193–1213.

Vaidya G, Lohman DJ, Meier R. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* **27:** 171–180.

Vollenweider P, Menard T, Arend M, Kuster TM, Gunthardt-Goerg MS. 2016. Structural changes associated with drought stress symptoms in foliage of Central European oaks. *Trees-Structure and Function* **30:** 883–900.

Vonhof H, Kaandorp R. 2010. Climate variation in Amazonia during the Neogene and the Quaternary. In: Hoorn C, Wesselingh F eds. *Amazonia, landscape, and species evolution: a look into the past.* Oxford: Wiley-Blackwell, 201–210.

Vonhof H, Wesselingh F, Ganssen G. 1998. Reconstruction of the Miocene western Amazonian aquatic system using molluscan isotopic signatures. *Palaeogeography Palaeoclimatology Palaeoecology* **141:** 85–93.

Warren DL, Glor RE, Turelli M. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* **62:** 2868–2883.

Warren DL, Glor RE, Turelli M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33:** 607–611.

Wesselingh F, Salo J. 2006. A Miocene perspective on the evolution of the Amazonian biota. *Scripta Geologica* **133:** 439–458.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig S1. Maximum likelihood tree reconstructed in RAxML for 57 *Astrocaryum* section *Huicungo* individuals within 13 species plus 22 outgroup within 6 species. Branch support SH-aLRT along each node. Individuals distributed in the Northern Amazonian Foreland Basin and the Central Andean Foothills are marked with branch colors red and blue, respectively. Bar represents number of substitutions per site.

Fig S2. Boxplots of A. precipitation of the driest month (bio14), B. precipitation seasonality (bio15), C. temperature of driest quarter (bio9), and D. elevation, derived from occurrence records for *Astrocaryum* section *Huicungo*. Boxes represent the interquartile range, horizontal lines within the boxes represent means and whiskers extend to the 95% most extreme data points.

Fig S3. 50% majority rule consensus tree resulting from a maximum parsimony analysis in PAUP* 4.0b10a for 14 *Astrocaryum* section *Huicungo* species and 12 outgroup based on the polymorphism of 35 anatomical and 31 morphological characters. Numbers along branches are bootstrap support. Species distributed in the Northern Amazonian Foreland Basin and Central Andean Foothills are marked in red and blue, respectively.

Table S1. Study species with information on sample identity in the DNA bank at the Institut de Recherche pour le Development (IRD) in France, herbarium voucher number, and Genbank (NCBI) accession number used to reconstruct the phylogenetic tree of *Astrocaryum* section *Huicungo*. Sequences in bold were newly generated for this study. Within the column "Chloroplast", Genbank numbers correspond to the following DNA sequences: *psbM-trnD, rps16* intron, *rps16-trnQ, trnT-trnD, trnG* intron, otherwise they correspond to plastid haplotypes obtained from Roncal *et al.* 2015. DNA region acronyms described in the methods section.

Table S2. List of 35 leaf anatomical and 31 morphological multistate characters obtained from the literature but newly or recoded in this study (Millán & Kahn, 2010; Kahn & Millán, 2013). Anatomical and morphological characters are numbered with an A and M, respectively.

Table S3. Data matrix of 31 morphological and 35 anatomical characters coded in this study. Multiple states within a species are indicated by a comma (polymorphic species). ?=missing data

Table S4. List of the 10 environmental variables used for the MAXENT distribution modeling of the Northern Amazonian Foreland Basin (NAFB) and Central Andean Foothills (CAF) clades. This table shows relative contributions of each environmental variable to each model as estimated by permutation importance values and Jackknife test gain.