The Neogene rise of the tropical Andes facilitated diversification of wax palms (*Ceroxylon*: Arecaceae) through geographical colonization and climatic niche separation

MARÍA J. SANÍN¹*, W. DANIEL KISSLING², CHRISTINE D. BACON³, FINN BORCHSENIUS⁴, GLORIA GALEANO⁵, JENS-CHRISTIAN SVENNING⁶, JOSÉ OLIVERA⁷, RINA RAMÍREZ⁷, PHILIP TRENEL⁸ and JEAN-CHRISTOPHE PINTAUD⁹†

¹Department of Biology, Universidad CES, Medellín, Colombia
²Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, Amsterdam, the Netherlands
³Department of Biological and Environmental Sciences, University of Gothenburg, Gothenburg, Sweden
⁴Science Museums, Aarhus University, Ole Worms Allé 1, Building 1137, DK-8000, Aarhus C, Denmark
⁵Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Bogota, Colombia
⁶Section for Ecoinformatics and Biodiversity, Department of Bioscience, Aarhus University, Aarhus, Denmark
⁷Universidad Nacional Mayor de San Marcos, Lima, Peru
⁸AgroTech a/S, Aarhus, Denmark
⁹Institut de Recherche pour le Développement, Montpellier, France

Received 4 November 2015; revised 20 February 2016; accepted for publication 27 February 2016

The tropical Andes are a biodiversity hotspot, partly due to their rich and complex floristic composition. A fundamental question regarding this outstanding biodiversity is what role the Andean orogeny has played in species diversification. *Ceroxylon* is a genus of endemic Andean palms that stands out in the palm family (Arecaceae) due to its adaptation to cold, mountainous environments. Here, we reconstruct the biogeography and climatic preference of this lineage to test the hypothesis that Andean uplift allowed diversification by providing suitable habitats along climatic and elevational gradients. Ancestral areas were reconstructed under a model allowing for founder-event speciation and climatic niches were inferred from climatic variables at present-day occurrences of all species. Niche evolution in a phylogenetic framework was evaluated by testing differences between the climatic niches of clades. Our analyses identified four main clades, with a general pattern of diversification through geographical colonization from south to north after the Pliocene uplift of the northern Andes. Adaptation to low temperatures was conserved at the generic level, with climatic niche differentiation among clades along elevational temperature gradients. We conclude that the Neogene Andean uplift has facilitated the diversification of this iconic plant group via opportunities for geographical migration and separation within its climatic niche. © 2016 The Linnean Society of London, Botanical Journal of the Linnean Society, 2016, 182, 303–317


INTRODUCTION

The tropical Andes is the most species-rich biodiversity hotspot on a global scale harbouring almost
twice as many plant species as the second-most diverse hotspot (Mesoamerica; Myers et al., 2000). Biotic composition within this region is complex and unifies diverse origins, including elements from Amazonia (e.g. Geonoma plants: Roncal et al., 2011), Central America (e.g. Chlorospingus birds: Weir et al., 2008; Hedysosmum plants: Antonelli & Sanmartin, 2011a), Guiana (e.g. Centroleineae frogs: Castroviejo-Fisher et al., 2009), the Holarctic (e.g. Lupinus plants: Hughes & Eastwood, 2006; Valeriana plants: Bell, Kutschker & Arroyo, 2012) and widespread lineages from across the Southern Hemisphere (e.g. Ceroxyleae plants: Trénel et al., 2007). Fundamental questions regarding this outstanding diversity are how and why it formed (Antonelli & Sanmartin, 2011b) and what role Andean orogeny (Neogene; 23–2.6 Mya) has played in the diversification and community assembly in the region. Palaeo-elevational estimates for the central and northern Andes are not yet fully understood, but an approximate chronology can be extracted from the general patterns of uplift from different sources of evidence (Gregory-Wodzicki, 2000). The central Andes reached higher elevations earlier in time than the northern Andes (e.g. Garzino et al., 2008). The Eastern Cordillera of Bolivia had attained 1000–1500 m during the middle Eocene (3500 m modern elevation), whereas the Eastern Cordillera of Colombia, connecting the Andes of Venezuela to the general course of the mountain chain, had attained only 0–500 m during this time and only 1000 m at c. 5 Mya (3165 m modern elevation; Gregory-Wodzicki, 2000). The central Andean plateau had attained elevations of c. 2000 m by 8 Mya, but present-day surface elevations over large areas were only reached during or after the Late Miocene (Mora et al., 2010).

A key issue regarding Andean orogeny is whether species have evolved and diversified in situ as a function of mountain orogeny, or if species dispersed into newly formed montane habitats made available only after surface uplift. Following this scenario, two potentially complementary hypotheses can be posited for the diversification of Andean lineages: (1) surface uplift has allowed in situ climate niche evolution through adaptation to climatic and environmental changes (niche shifts), and (2) colonization of newly available habitats created by orogeny was possible through migration while maintaining the ancestral niche (‘niche conservatism’, e.g. Donoghue, 2008). These two hypotheses are not necessarily exclusive but can be tested by the examination of climatic niche changes along the phylogenies of Andean species.

Niche evolution can be tested by reconstructing the niche of different clades in a dated phylogeny (e.g. Litsios et al., 2012) or by linking phylogenetic information with distributional data and environmental layers (e.g. Graham et al., 2004), for example climatic information linked to localities to quantify the climatic niche (e.g. Schnitzler et al., 2012). This realized niche describes an integral aspect of the environment of species (Pearman et al., 2008). For example, climatic variables such as temperature and precipitation have been identified as explanatory variables for the ecological distribution for several Ecuadorian Dentrobatid frogs (Graham et al., 2004). If niches are conserved among related taxa within a phylogeny, few differences in the climatic niche described by the climatic variables are expected across closely related clades. On the other hand, if the niche expands or shifts along the phylogeny, then differences in climatic niche axes should be detected between closely related taxa.

Phylogenetic niche conservatism describes the tendency for related species to have similar niches (Harvey & Pagel, 1991) and is considered to be a primary contributor to large-scale patterns of biodiversity (e.g. Pyron et al., 2015). For instance, if a lineage is adapted to certain climatic conditions, then descendants of that lineage are likely to remain adapted for a similar niche, as has been shown for arctic plants (Tkach et al., 2014) and North American forest communities (Hawkins et al., 2014). Despite strong evidence for niche conservatism across the tree of life (e.g. Crisp et al., 2009; Wiens et al., 2010; Eisenhardt et al., 2015), some studies have shown no positive correlation between phylogenetic distance and niche differentiation, particularly in isolated conditions such as on islands (e.g. plants in the Canary Islands: Kim et al., 1996; Anolis lizards: Knouft et al., 2006; Losos et al., 2003) or in high-elevation forest patches (e.g. Aphonoloma birds: Rice, Martínez-Meyer & Peterson, 2003). Without phylogenetic niche conservatism, related species expand their niches or experience niche shifts primarily through interspecific competition, which causes segregation along ecological axes and ultimately speciation (Silvertown, 2004). In cases of little coexistence between sister taxa, there are no limitations for phylogenetic conservatism in terms of biotic interactions (Webb et al., 2002).

With >2500 species, Arecaceae (palms) are primarily distributed in tropical regions (Dransfield et al., 2008), making them an excellent model system for studying tropical forest evolution (Bacon, 2013; Couvreur & Baker, 2013). The Neotropics harbours extraordinary palm species diversity, suggested to be the result of high in situ diversification, long-term geographical isolation and relatively stable climatic conditions during the Cenozoic (Kissling et al., 2012a). Relatively few palm species are distributed in cooler subtropical or alpine environments.
in the genus Ceroxylon is poorly known [see Trénel et al., 2007], the interspecific relationships among the species of this genus show niche differentiation along temperature and precipitation gradients by linking localities and climatic data (i.e. annual elevation and precipitation). This suggests that understanding the role played by climatic adaptation in the diversification of the genus is crucial. One of the few cold-adapted palm genera is Ceroxylon Bonpl. (the wax palms). Ceroxylon and Parajubaea Burret are the only two palm genera that are endemic to Andean alpine forests and Cer- oxylon includes the palm species that are found at the highest elevations in the world.

Ceroxylon encompasses 12 currently accepted species that grow in moist cloud forests at 800–3500 m elevation, most commonly at ~2000 m (Sanín & Galeano, 2011). The genus is one of four in tribe Ceroxyleae, the others being the most species-rich genus in the tribe, Ravenea H.Wendl. ex C.D.Bouché (> 20 species) from Madagascar and the Comoro Islands, and two monotypic genera, Juania Drude from the Juan Fernández Archipelago and Oraniopsis J.Dransf., A.K.Irvine & N.W.Uhl from Australia. In a previous study (Trénel et al., 2007), Eocene dispersal through Antarctica was shown to be the driver of the southern geographical disjunction of Ceroxyleae, rejecting the alternative hypothesis that it was ancient Gondwanan distribution subsequently divided by vicariance. Although a robust phylogenetic hypothesis exists for Ceroxyleae at the generic level, placing Ceroxylon as sister to the oceanic, sub-temperate Juania (Trénel et al., 2007), the interspecific relationships in the genus Ceroxylon, which are probably key to understanding the role played by climatic adaptation to mountain habitats in Andean evolution, remain poorly known [see Trénel et al. (2007) and Baker et al. (2009) for previous phylogenetic analyses].

Here, we reconstruct a time-calibrated phylogenetic tree with samples of all species, aiming at clarifying species relationships in Ceroxylon. We also reconstruct the biogeographical history and niche evolution in this cold-hardy genus. Because adaptation to cold environments is rare in palms, we predict that Ceroxylon was pre-adapted to cold conditions and diversified after geographical colonization from south to north by conserving an inherited climatic niche of adaptation to cold (niche conservatism), rather than expanding earlier and then shifting in situ to this niche after Andean uplift. We further test to what extent clades in the genus show niche differentiation along temperature (elevation) and precipitation gradients by linking species locality records to climatic data (i.e. annual mean temperature, precipitation, seasonality, extremes of drought and cold). Our analysis makes a step forward by improving the phylogenetic resolution in Ceroxylon. It also contributes to a deeper understanding of the complex diversification history in the Andean biodiversity hotspot by demonstrating that speciation may respond to segregation of the climatic niche along elevational temperature gradients, especially in cold-adapted lineages.

MATERIAL AND METHODS

Taxon sampling

Samples of 20 individuals were sequenced, including one or two samples of each of the 12 Ceroxylon spp. accepted in the monograph of the genus (Sanín & Galeano, 2011). One sample of each of the three other genera of Ceroxyleae were also included: Juania australis (Mart.) Drude ex Hook.f., Oraniopsis appendiculata (F.M.Bailey) J.Dransf., A.K.Irvine & N.W.Uhl and Ravenea xerophila Jum. (Trénel et al., 2007). One representative of each of the other palm subfamilies were used as outgroups: Astrocaryum murumuru Mart. (Arecoideae), Mauritia flexuosa L.f. (Calamoideae), Phoenix dactylifera L. (Coryphoideae) and Nypa fruticans Wurmb (Nypoideae). Data for the nuclear low-copy gene PRK (Roncal et al., 2005) and the plastid coding locus ndhF (Cuenca & Asmsen-Lange, 2007) were generated to allow for expanded sampling from a previous study (Trénel et al., 2007). Additionally, new datasets for the plastid loci atpF-atpH and atpH-atpI were generated following Scarcelli et al. (2011). Sequences were aligned using default parameters in the Muscle algorithm (Edgar, 2004) in Geneious 5.1 (Kearse et al., 2012). The alignments obtained were then exported and manually refined following Simmons (2004) in Mesquite (Maddison & Maddison, 2015). Specimens, voucher information, localities and GenBank accession numbers are listed in Table 1.

Phylogenetic analysis and divergence time estimation

Due to the inherent linkage of plastid DNA loci, they were treated as a single locus and analysed as one partition separate from the nuclear gene in BEAST v.1.7.5 (Drummond & Rambaut, 2007). The data were partitioned by locus to allow for variation in substitution models, which can alleviate or remove the impact of saturation caused by using relatively older calibrations for inferences directed at more recently evolved nodes (sensu Ho et al., 2005). Substitution model selection was done in MrModelTest implemented in Mega 6.06 (Tamura et al., 2013). The analysis was run using an uncorrelated lognormal molecular clock model, a Yule pure birth speciation tree model with no starting tree model (without specifying a starting tree) and the GTR + Γ model of nucleotide substitution with four rate categories. The following species were
Table 1. Genbank accession numbers and voucher information of sequences used, the numbers with a shaded background were newly generated for this study.

<table>
<thead>
<tr>
<th>Species</th>
<th>DNA code</th>
<th>Country</th>
<th>Herbarium voucher</th>
<th>ndhF</th>
<th>atpF-atpH</th>
<th>atpH-atpI</th>
<th>PRK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phoenix dactylifera</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GU811709</td>
<td>GU811709</td>
<td>GU811709</td>
<td>-</td>
</tr>
<tr>
<td>Mauritia flexuosa</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>EU186209</td>
<td>-</td>
<td>-</td>
<td>KJ500131</td>
</tr>
<tr>
<td>Nypa fruticans</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>EUR8217</td>
<td>-</td>
<td>-</td>
<td>AJ831357</td>
</tr>
<tr>
<td>Astrocaryum murumuru</td>
<td>JCP 0905</td>
<td>French Guiana</td>
<td>A. Perez 718 (CAV)</td>
<td>EU004593</td>
<td>KP255421</td>
<td>KP339836</td>
<td>HQ265590</td>
</tr>
<tr>
<td>Ravena xerophila</td>
<td>JCP 2415</td>
<td>French Guiana</td>
<td>P. Trenel 008 (AAU)</td>
<td>EF128298</td>
<td>KP255440</td>
<td>KP339855</td>
<td>EF128382</td>
</tr>
<tr>
<td>Orianopsis appendiculata</td>
<td>JCP 2416</td>
<td>French Guiana</td>
<td>1988-227 (K)</td>
<td>EF128300</td>
<td>KP255441</td>
<td>KP339856</td>
<td>AJ831359</td>
</tr>
<tr>
<td>Juania australis</td>
<td>JCP 1454/2412</td>
<td>Chile</td>
<td>P. Trenel 004 (AAU)</td>
<td>EF128299</td>
<td>KP255439</td>
<td>KP339854</td>
<td>EF128383</td>
</tr>
<tr>
<td>Ceroxylon pityrophyllum</td>
<td>JCP 2378</td>
<td>Bolivia</td>
<td>Nee 30324 (LPB)</td>
<td>KM669742</td>
<td>KP255431</td>
<td>KP339846</td>
<td>KP268627</td>
</tr>
<tr>
<td>Ceroxylon pityrophyllum</td>
<td>JCP 1512/1514</td>
<td>Peru</td>
<td>Kahn 4478 (USM)</td>
<td>KM669735</td>
<td>KP255422</td>
<td>KP339837</td>
<td>KP268621</td>
</tr>
<tr>
<td>Ceroxylon vogelianum</td>
<td>JCP 2366</td>
<td>Peru</td>
<td>Borchsenius 640 (AAU)</td>
<td>KM669737</td>
<td>KP255426</td>
<td>KP339841</td>
<td>EF128396</td>
</tr>
<tr>
<td>Ceroxylon vogelianum</td>
<td>JCP 2332</td>
<td>Colombia</td>
<td>Bernal 2475 (COL)</td>
<td>KM669743</td>
<td>KP255432</td>
<td>KP339847</td>
<td>KP268634</td>
</tr>
<tr>
<td>Ceroxylon alpinum</td>
<td>JCP 2367</td>
<td>Colombia</td>
<td>Rodriguez s.n. (COL)</td>
<td>KM669738</td>
<td>KP255427</td>
<td>KP339842</td>
<td>KP268623</td>
</tr>
<tr>
<td>Ceroxylon alpinum</td>
<td>JCP 2371</td>
<td>Colombia</td>
<td>Bernal 2320 (COL)</td>
<td>KM669739</td>
<td>KP255428</td>
<td>KP339843</td>
<td>KP268624</td>
</tr>
<tr>
<td>Ceroxylon parafirans</td>
<td>JCP 2357</td>
<td>Ecuador</td>
<td>Balslev 62553 (AAU)</td>
<td>EF128308</td>
<td>KP255423</td>
<td>KP339838</td>
<td>EF128391</td>
</tr>
<tr>
<td>Ceroxylon parafirans</td>
<td>JCP 2404</td>
<td>Colombia</td>
<td>Sanin 017 (COL)</td>
<td>KM669749</td>
<td>KP255438</td>
<td>KP339853</td>
<td>KP268622</td>
</tr>
<tr>
<td>Ceroxylon parvum</td>
<td>JCP 2359</td>
<td>Ecuador</td>
<td>Borchsenius 649 (AAU)</td>
<td>EF128310</td>
<td>KP255424</td>
<td>KP339839</td>
<td>EF128392</td>
</tr>
<tr>
<td>Ceroxylon amazunicum</td>
<td>JCP 2422</td>
<td>Ecuador</td>
<td>Trenel 017 (AAU)</td>
<td>KM669750</td>
<td>KP255442</td>
<td>KP339857</td>
<td>KP268633</td>
</tr>
<tr>
<td>Ceroxylon amazunicum</td>
<td>JCP 2373</td>
<td>Ecuador</td>
<td>Balslev 6452 (AAU)</td>
<td>KM669741</td>
<td>KP255430</td>
<td>KP339845</td>
<td>KP268626</td>
</tr>
<tr>
<td>Ceroxylon sasaimae</td>
<td>JCP 2372</td>
<td>Colombia</td>
<td>Bernal 2311 (COL)</td>
<td>KM669740</td>
<td>KP255429</td>
<td>KP339844</td>
<td>KP268625</td>
</tr>
<tr>
<td>Ceroxylon peruvianum</td>
<td>JCP 2394</td>
<td>Colombia</td>
<td>Sanin 005 (COL)</td>
<td>KM669746</td>
<td>KP255435</td>
<td>KP339850</td>
<td>KP268629</td>
</tr>
<tr>
<td>Ceroxylon echinulatum</td>
<td>AU821</td>
<td>Ecuador</td>
<td>Borchsenius 647 (AAU)</td>
<td>EF128305</td>
<td>KP255418</td>
<td>KP339833</td>
<td>EF128388</td>
</tr>
<tr>
<td>Ceroxylon echinulatum</td>
<td>AU995</td>
<td>Ecuador</td>
<td>Borchsenius 639 (AAU)</td>
<td>EF128306</td>
<td>KP255419</td>
<td>KP339834</td>
<td>EF128389</td>
</tr>
<tr>
<td>Ceroxylon ventricosum</td>
<td>JCP 2356</td>
<td>French Guiana</td>
<td>Borchsenius 646 (AAU)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>EF128394</td>
</tr>
<tr>
<td>Ceroxylon ventricosum</td>
<td>JCP 2387</td>
<td>Ecuador</td>
<td>Trenel s.n. (AAU)</td>
<td>KM669744</td>
<td>KP255433</td>
<td>KP339848</td>
<td>KP268635</td>
</tr>
<tr>
<td>Ceroxylon ventricosum</td>
<td>JCP 2402</td>
<td>Colombia</td>
<td>Sanin 015 (COL)</td>
<td>KM669748</td>
<td>KP255437</td>
<td>KP339852</td>
<td>KP268631</td>
</tr>
<tr>
<td>Ceroxylon quindiuense</td>
<td>JCP 2392</td>
<td>Peru</td>
<td>Sanin 003 (COL)</td>
<td>KM669745</td>
<td>KP255434</td>
<td>KP339849</td>
<td>KP268628</td>
</tr>
<tr>
<td>Ceroxylon quindiuense</td>
<td>JCP 2397</td>
<td>Colombia</td>
<td>Sanin 008 (COL)</td>
<td>KM669747</td>
<td>KP255436</td>
<td>KP339851</td>
<td>KP268630</td>
</tr>
<tr>
<td>Ceroxylon ceriferum</td>
<td>JCP 2360</td>
<td>Colombia</td>
<td>Galeano 7584 (COL)</td>
<td>KM669736</td>
<td>KP255425</td>
<td>KP339840</td>
<td>KP268622</td>
</tr>
<tr>
<td>Ceroxylon ceriferum</td>
<td>cer05</td>
<td>Colombia</td>
<td>Galeano 7584 (COL)</td>
<td>KM669751</td>
<td>KP255420</td>
<td>KP339835</td>
<td>KP268620</td>
</tr>
</tbody>
</table>
We used ancestral area reconstruction to test the geographical colonization of *C. amazonicum* Galeano, *C. ceriferum* (H.Karst.) Pittier, *C. echinulatum* Galeano, *C. parvifrons* (Engel) H.Wendl., *C. ptyropolyphylum* (Mart.) Mart. ex H.Wendl. and *C. vogelianum* (Engel) H.Wendl. The Markov chains were run for 50 million generations, sampling every 1000 generations, and repeated three times to test for Markov chain Monte Carlo chain convergence and to ensure that effective sample sizes exceeded 200. After verifying that chains had reached stationarity in Tracer v.1.5, 50% of trees were removed as burn-in and BEAST log files were combined in LogCombiner v.1.5.4. Tree files were combined to estimate median node height and the 95% highest posterior density (HPD) in TreeAnnotator v.1.5.4.

The phylogeny was co-estimated with divergence times. Because we do not know of any unambiguously identified fossils in Ceroxyloideae, we used two secondary calibration points obtained from a broader study based on primary fossil calibrations in Arecaceae (Couvreur, Forest & Baker, 2011). A normal distribution for the secondary calibration point at the root node of the phylogenetic tree (the crown of Arecaceae) was implemented using a mean of 100 Mya and a 6.0 SD of the prior to capture the variation around the mean (109–92 Mya Arecaceae crown age; Couvreur *et al.*, 2011). A normal distribution was also used to constrain the crown of Ceroxyleae at 17 Mya (29–7 Mya HPD; using a mean of the distribution at 17 and a SD of 6). The normal distribution has been shown to be most appropriate for modelling secondary calibrations because it reflects the uncertainty in secondary date estimates (Ho, 2007; Ho & Phillips, 2009).

**Ancestral area reconstruction**

We used ancestral area reconstruction to test the geographical pattern of diversification history, i.e. whether geographical colonization of *Ceroxylon* of the Andes through a south-to-north expansion is supported. Biogeographical regions were defined by geographical barriers in the Andes following Mora *et al.* (2010) and taking into account climatic availability in the Neogene time frame of this study: Cordillera de la Costa (northern Venezuela; area A); Sierra Nevada de Santa Marta (Colombia; area B); Merida Cordillera (Táchira to Lara states, Venezuela; area C); Colombian Cordilleras (area D); western slopes of Ecuador (area E); Peru north of the Abancay deflection and eastern slopes of Ecuador (area F); south of the Abancay deflection and north of Bolivian orocline (Bolivia and Peru; area G); and the tropical Andes south of the Bolivian orocline (area H).

To reconstruct ancestral areas we used the package BioGeoBEARS v.0.2.1 (Matzke, 2013, 2014) in R (R Development Core Team, 2015). BioGeoBEARS implements widely used models of range evolution [e.g. dispersal extinction cladogenesis (DEC); Ree & Smith, 2008] and includes an additional jump parameter ‘*J*’ that corresponds to cladogenetic speciation mediated by founder events (Matzke, 2014). This parameter allows daughter species to ‘jump’ outside the geographical range of parental species, as is highly appropriate for Andean lineages that colonize new niches or experience niche shifts. To reconstruct ancestral ranges we focused strictly on *Ceroxylon*, excluding the other taxa in the tree due to their non-Andean distribution. We computed the DEC model with the J parameter using the maximum clade credibility tree inferred in BEAST. We constrained the number of ancestral areas to six to correspond to the widest distribution present in the extant areas (Ronquist, 1997; Matzke, 2014). We also constrained the presence of ancestral species in areas and dispersal between areas during the following time slices: (1) 8.3–5.6 Mya: *Ceroxylon* not present in areas A, B, C or D and zero probability of dispersal between areas G and F and between F and E; (2) between 5.6 and 4.5 Mya: zero probability of dispersal to areas A, C and D; and (3) between 4.5 and 4.3 Mya: zero probability of dispersal from area D to areas A or C. The time-slice constraints incorporate geological evidence of existence of these areas into the models based on Mora *et al.* (2010) and Gregory-Wodzicki (2000).

**Climatic niche evolution**

To examine climatic niche evolution, 247 occurrence records (i.e. latitude and longitude coordinates of *Ceroxylon* spp.) were collected from 19 herbaria (herbarium codes: AAU, B, BH, COL, HUA, JBSC, K, LE, LPB, MER, MERC, MO, NY, P, QCA, QCNE, SCZ, US and USM), derived from our fieldwork in Bolivia, Colombia, Ecuador, Peru and Venezuela and from the Global Biodiversity Information Facility (GBIF). All records were quality-checked and reviewed for geographical and taxonomic issues resulting in the correction or exclusion of erroneous georeferences and/or individuals with reference to the monograph of *Ceroxylon* by Sanín & Galeano (2011).

We extracted information on climatic niches using occurrence records in ArcGIS with the 19 climatic variables from the WorldClim dataset (Hijmans *et al.*, 2005). We separated the WorldClim data into temperature-related (WorldClim Bio1–Bio11) and precipitation-related variables (WorldClim Bio12–Bio19). We further extracted the elevation of each occurrence record by intersecting them with elevational data from the Shuttle Radar Topography Mission (SRTM) elevation database (aggregated to 30 arc-seconds, ‘1 km’) (available at http://www.world
clim.org and http://www2.jpl.nasa.gov/srtm/). Although other environmental factors (e.g. soil) may influence species distributions, we focused on climatic and elevational parameters because they are the key drivers of large-scale distributions of palms (Tomlinson, 2006; Kissling et al., 2012b), they can be obtained from public databases and because the steep Andean climatic gradients probably have a strong impact on clade distributions and evolutionary dynamics at this spatial scale (Parra, Graham & Freile, 2004). To make climatic niche comparisons more robust and given that for some species the occurrence records are scarce (i.e. some species are endemic to small areas), we analysed the climatic data for clades instead of species, based on the clades reconstructed in the phylogenetic tree (a number was assigned to each clade).

We performed separate principal components analyses (PCAs) for each set of climate data (i.e. temperature- and precipitation-related variables) and extracted the first two PCA axes to describe the main climatic gradients for both. We further used the elevational data from the occurrence records to quantify the elevational distribution of clades. Climate data were log-transformed and elevational data remained untransformed. We then performed the Tukey honest significant distance (HSD) test, which is a parametric test and allows for multiple pairwise comparisons of means such as PCA climate niche parameters and elevational distributions among the four Ceroxylon clades. We used a kernel density function for the climate data to visualize the distribution of clades within environmental space according to the PCA results. If the climatic niche parameters (i.e. mean PCA values) of the clades were not statistically different from that of the earliest divergent clade, phylogenetic niche conservatism was assumed and if parameters differed, niche shifts were assumed. We further assumed niche differentiation if climatic niche parameters differed between sister clades in Ceroxylon. As a coarse comparison with Ceroxylon, we extracted data on annual temperature (Bio1) from occurrence points of Juania australis. As J. australis has a restricted distribution (i.e. all spatial occurrences are aggregated on a 93-km² island), other climatic variables were not extracted. All statistical analyses of climate data were performed with the R software platform (R Core Team, 2015), packages foreign (R Core Team, 2015) and MASS (Venables & Ripley, 2002).

RESULTS

PHYLOGENETIC AND DIVergence TIME ANALYSES

Four principal clades were recovered in Ceroxylon (Figs 1, 2; Supporting Information, Fig. S1): clade 1 representing the early divergent sister species pair C. pityrophyllum and C. vogelianum [1.00 posterior probability (PP)]; clade 2 composed of all the mid-elevation species C. alpinum Bonpl. ex DC., C. parvifrons Galeano, C. sasaimae Galeano, C. echinulatum, C. peruviano Galeano, Sanín & K.Mejia and C. amazonicum (1.00 PP); clade 3 composed of only the widely distributed C. parvifrons (1.00 PP); and clade 4 composed of the three upper-elevation species C. ventricosum Burret, C. quindiuense (H.Karst.) H.Wendl. and C. ceraferum (0.50 PP). The use of two normally distributed calibrations yielded high effective sample size values (>1000) and consistent results between Markov chain Monte Carlo chain iterations. The stem and crown nodes of Ceroxylon (Supporting Information, Fig. S1; Fig. 2) were inferred at 12.3 Mya (16.8–8.3 Mya HPD) and 8.3 Mya (11.6–5.3 Mya HPD), respectively. Each of the major clades resolved in the genus diversified concurrently, with the mean crown ages being 2.6, 3.7, 0.4 and 4.5 Mya for clades 1–4, respectively.

ANCESTRAL AREA RECONSTRUCTION

The inferred ancestral distribution for Ceroxylon and its ancestral nodes are to some extent ambiguous (indicated in white on the pie charts in Fig. 2), probably due to the wide ranges of C. parvifrons and C. vogelianum. The most probable ancestral area is the eastern part of the Ecuadorian Andes and central Peru (area F). The ancestral distribution of clade 1 is largely ambiguous, but subsequent divergence leads to two clades inferred in areas GH (southern Peru and Bolivia) and CDEF (from Venezuela to Peru), respectively (Fig. 2). The most likely distribution for clade 2 was also area F. Clade 3 was inferred to have a wide Andean distribution (areas A, C, E, D, F and G in Fig. 2), excluding the extreme southern portion and the northern Sierra Nevada de Santa Marta (Bolivia and Colombia, respectively). Clade 4 was also inferred to have an ancestral distribution in area F, with subsequent disjunction of two clades, one in F and one in adjacent area D (Continental Cordilleras), representing a shift from the Andes of Peru and Ecuador northwards into Colombia. This clade in D also subsequently diverges into two nodes, one in D and one in the coastal mountains of Venezuela and Colombia (areas A–B), further north.

Niche Reconstruction

In the PCAs of climatic factors, 89.6% of the variability in temperature-related data and 94.9% in precipitation-related data were explained by axes 1 and 2 of the respective PCAs. The first axes of both variable
sets were largely related to average annual variables and the second axes related to seasonality. The four clades showed the most marked differentiation with respect to the annual temperature PCA axis. All clade pairs showed statistically significant differences for annual temperature PCA values, except for clades 1 and 4 (Fig. 3; Supporting Information, Table S1). Among Ceroxylon clades, a group including clades 2, 3 and 4 had different temperature seasonality PCA values than clade 1. Sister clades 2 and 3 had different annual temperature and precipitation PCA values. Clades 2 and 3 were not statistically different from clades 1 and 4 in the annual precipitation PCA axis, suggesting that this aspect of the climatic niche has differentiated within rather than among the clades (Fig. 3; Supporting Information, Table S1). The annual temperature recovered for Juania australis (13.6 °C) is found at the lower range of the temperatures covered by Ceroxylon clades (17 ± 3.4 °C). The precipitation seasonality PCA axis did not statistically differ among clades. Hence, a general pattern of niche similarity was evident between Ceroxylon and Juania and was even conserved through the early divergent Ceroxylon clade 1. Furthermore, there were significant differences between sister clades 2 and 3, suggesting niche differentiation in temperature-related variables of the climatic space (Fig. 3). This difference in climatic
Figure 2. Maximum clade credibility tree of *Ceroxylon* and inferred ancestral areas. Pie charts represent the probability of an area and area delimitation is depicted in the inset map (lower left). The white sections of the pie charts represent all distributions with a probability of < 0.5 inferred for that node and may represent more than one area. Mean node ages (left of pie charts, in bold type) and the posterior probabilities (PP) of branch support (right of pie charts) for major clades are listed at the nodes. All PP values are 1.0, unless otherwise listed. Four main clades are identified in the phylogenetic tree (colour bars on the right-hand side with clade numbers as used in Figure 1, and in the Discussion, respectively).
Figure 3. Climatic niche space of the four Ceroxylon clades (1–4 from left to right, with same colors as in Figs 1, 2). Top and centre rows: climatic niches were derived by intersecting species occurrence records with 19 climate variables (WorldClim). Climate data were grouped into two datasets (temperature: WorldClim Bio1–Bio11; precipitation: WorldClim Bio12–Bio19) and then subjected to separate principal components analysis (PCA). The PCA axes represent annual and seasonality values for temperature and precipitation. Contour lines indicate kernel density estimates for locality records of a given clade, while grey points indicate the other clades for reference. Bottom row: box plots of climatic variables for Ceroxylon clades 1–4 (x-axis). Boxes represent the interquartile range, horizontal lines within the boxes represent medians and whiskers extend to the 95% most extreme data points. Values on the y-axis correspond to those of the PCA and correlate with temperature and precipitation annual averages and seasonality. Letters above whiskers indicate statistical differences ($P > 0.05$) between pairs of clades for each axis of the PCA through multiple pairwise comparisons of means with the Tukey HSD test (compare Supporting Information, Table S1).
niches between clades 2 and 3 corresponded mainly to geographical segregation along the elevational gradient (Fig. 1E; Supporting Information, Table S2), i.e. to clade specialization on a specific area of the elevational range covered by the genus.

**DISCUSSION**

**PHYLOGENETIC RELATIONSHIPS**

The phylogenetic results presented here are a significant advance in our understanding of the species relationships in *Ceroxylon* and are based on complete species-level sampling, providing a solid basis for biogeographical and ecological inference. Our results support a Pliocene colonization of the northern Andes following mountain uplift and we suggest this to be the main condition for diversification in *Ceroxylon*. The clade-level resolution obtained in the phylogenetic analysis is a significant advance compared to previous phylogenetic trees in which infrageneric relationships were not supported (Trénél et al., 2007; Baker et al., 2009). This allowed for testing hypotheses regarding climatic niche evolution in the genus.

Although the present analysis provides substantial internal resolution in *Ceroxylon*, the low resolution and support within clades is probably due to their recent origin (4.5–0.4 Mya; Fig. 2) and the limited variation in the DNA sequence markers. Given the long generation times of wax palms (c. 83 years: Sanín et al., 2013), mutation and genetic drift are likely to occur at slow rates, extending the time required for lineages to sort allelic variation (Weller & Wu, 2015). Nevertheless, we identified four clades in the genus. A detailed description of the recovered clades, in terms of the morphological characters, the divergence times and distributions of the species recovered in each clade, is provided below.

Clade 1 is composed of two species, *C. pityrophyllum* and *C. vogelianum*. Sanín & Galeano (2011) resurrected *C. pityrophyllum* and also considered *C. vogelianum* as an accepted species. In the analysis of Trénél et al. (2007), which pre-dated the taxonomic revision of Sanín & Galeano (2011), *C. pityrophyllum* was synonymized with *C. parvum* (clade 2), resulting in polyphyly of the species. *Ceroxylon vogelianum* and *C. pityrophyllum* have a similar habit, both have flower whors in parts of three (otherwise widespread in the tribe) and they are the only species in the genus with pollen exine that is not reticulate (Sanín & Galeano, 2011). These species represent a clear case of complete allopatry: *C. pityrophyllum* grows on the eastern slopes of the Bolivian and Peruvian Andes from Chuquisaca to Cuzco; and *C. vogelianum* north of Cuzco to the north through to Venezuela. Also, *C. pityrophyllum* has the southern-most distribution in the genus.

Clade 2 is the most species-rich (seven species) and species are inferred to have diverged only recently (c. 3.7 Mya; Fig. 2). For example, a previous study showed that *C. echinulatum* diverged by separation on the eastern and western slopes of the Andes, estimated to have happened at 0.6 Mya (Trénél et al., 2008). Several of these species are microendemics (species with highly restricted distributions), *C. amazonicum*, *C. parvum*, *C. peruvianum* and *C. sasaimae*, which may reflect their recent origin, potentially due to less time to disperse.

The three currently recognized species of clade 4 occupy the same elevation range, growing at 1900–3000–3150 m (based on field data; Sanín & Galeano, 2011). However, they replace each other latitudinally (Sanín & Galeano, 2011) and are thus never sympatric. *Ceroxylon ceriferum* is restricted to the Coastal Cordilleras at 9–11°N, whereas *C. quindiuense* and *C. ventricosum* grow at different latitudes along the Cordilleras of Colombia, Ecuador and Peru from 4°S to 8°N. *Ceroxylon ventricosum* and *C. quindiuense* are not recovered as monophyletic, which could be expected in such a recent group, due to incomplete lineage sorting (see Naciri & Linder, 2015). The distribution of *C. quindiuense* as currently recognized is disjunct with populations across the Colombian Andes and isolated populations in northern Peru (Galeano et al., 2008). In the geographical gap on both slopes of the Ecuadorian Andes and southern Colombia, the species *C. ventricosum* is found (Fig. 1D). Our phylogenetic tree suggests that individuals from Colombia circumscribed under *C. ventricosum* are more closely related to *C. quindiuense* from Colombia and that *C. quindiuense* from Peru and Colombia may be different entities. Further genetic data and analyses will be needed to elucidate the phylogenetic relationships in clade 4. *Ceroxylon ceriferum* is a well-supported species sister to paraphyletic *C. quindiuense* and *C. ventricosum* as currently circumscribed.

**SPECIES DIVERSIFICATION FOLLOWING ANDEAN UPLIFT FROM SOUTH TO NORTH**

*Juania* was reconstructed as the sister genus of *Ceroxylon* by Trénél et al. (2007) and in other major phylogenetic analyses of palms (Asmussen et al., 2006; Baker et al., 2009) and this relationship is also recovered here. The mean stem age of Neotropical *Ceroxyleae* (*Ceroxylon* + *Juania*) is estimated at 15.2 Mya (Supporting Information, Fig. S1). The divergence of *Ceroxylon* and *Juania* is inferred at 12.3 Mya, an age significantly older than Robinson
Crusoe Island where *Juania* is currently endemic (4 Mya; Stuessy et al., 1984). Dispersal to and successful colonization of old lineages on young islands is well documented (e.g. Garcia-Verdugo & Fay, 2014, and references therein) and these populations often persist longer than their source populations on the continent due to the ocean-buffered climate of islands (Heads, 2014). Ancestral lineages of *Ceroxylon* and *Juania* were probably affected by desertification in south-western South America (von Rotz et al., 2005). Desertification of the coast of Chile and southern Peru is estimated at 22.5 Mya (von Rotz et al., 2005), whereas major uplift of the Andes at 33°S, the latitude where *Juania* occurs today, was dated at 25 Mya (Gutierrez et al., 2013). These geographical changes potentially provided new suitable environment eastwards, where *Ceroxylon* could have differentiated from the *Juania* + *Ceroxylon* ancestor. Indeed, palaeoevidence of montane cloud forest flora, the only habitat of *Ceroxylon*, in the eastern cordillera of Bolivia from the Miocene (Graham, Gregory-Wodzicki & Wright, 2001) supports the idea. The north-eastern Andes had not attained cloud forest elevations at the estimated crown age of *Ceroxylon* (8.3 Mya), suggesting that the central Andes was the only area that could provide a suitable habitat for ancestral lineages under a climatic niche conservation scenario. Other cool habitats at subtropical latitudes were dry during this time window (e.g. Gregory-Wodzicki, 2007) and hence unsuitable for *Juania* + *Ceroxylon*.

The crown ages of clades 2–4 at 5.6 Mya are consistent with our hypothesis that initial diversification was linked to the main uplift pattern of the northern Andes (Fig. 2). At this time, the central Andes of Peru and Bolivia had attained much of their actual elevation (Gregory-Wodzicki, 2000). Similarly, splits in clades 2 and 4 are all younger than 5 Mya, coinciding with surface uplift in the northern Andes (Gregory-Wodzicki, 2000) and with the shortening of the Abancay deflection in the Late Miocene (Espurt et al., 2011).

The pattern of colonization is evident at the clade level. Clade 4 was present at 4.5 Mya in area F (Peru north of the Abancay deflection and eastern slopes of Ecuador) and split into daughter nodes occupying areas F and D (Colombian cordilleras), where the node in D eventually splits into daughter nodes in D and A–B (northern Venezuela and Sierra Nevada de Santa Marta in Colombia). These ancestral area inferences suggest a pattern of founder events of the adjacent areas to the north of the ancestral distribution. Clades 2 and 4 include lineages that colonized the northern Andes and that occupy different areas of climatic space (Fig. 3; Supporting Information, Table S1), indicating that colonization with speciation was possible through the elevational segregation of species.

The general biogeographical pattern of south-to-north migration following Andean uplift has been described for several other Andean plants such as *Oreobolus* R.Br. (Chacón et al., 2006), potatoes and their cyst-nematodes (Picard, Sempere & Plantard, 2007), *Perezia* Lag. (Simpson et al., 2009), Barnadesioideae (Gruensteaudl et al., 2009), Azorella Lam. (Nicolás & Plunkett, 2012), Bomarea (Chacón et al., 2012), *Puya* Molina (Jabaily & Sytsma, 2013), Solanaceae (Sárkinen et al., 2013) and Apiaceae (Nicolás & Plunkett, 2014) and in animals such as Proctoporus lizards (Doan, 2003) and Adelomyia hummingbirds (Chaves, Weir & Smith, 2011; see Luebert & Weigend, 2014, for a comprehensive review). Our results thus agree with the general pattern of Andean biogeography observed in other groups.

Niche conservatism vs. niche differentiation

We found the environmental niche of *Ceroxylon* to be mostly conserved, with orogeny driving diversification in the genus through elevational segregation of clades. Examination of the climatic niche of the genus provides evidence for the drivers of speciation in *Ceroxylon*. Temperature was found to be a more informative parameter than precipitation in differentiating the climatic niches of the four *Ceroxylon* clades (Fig. 3). Our results suggest that diversification in the genus is linked to climatic niche segregation of temperature-related variables. Various pairwise comparisons of clades show significant differences in the annual temperature niche (five out of six comparisons; Fig. 3 bottom row; Supporting Information, Table S1) and there is a clear reduction in the occupied range of temperature seasonality values in clades 2–4 as opposed to clade 1 (Fig. 3, top row). Taken together, these results show that clade 1 has an intermediate annual temperature niche and a wider temperature seasonality niche than the other three clades and that clades 2–4 have segregated the annual temperature space into specific portions. Clades 2 and 3 have differentiated their annual temperature niche by occupying a part of the niche of clade 1 (Fig. 3).

We did not detect a pattern of niche shifts and even if clades have extended into the extremes along the annual temperature axis, their occupied niches greatly overlap. Since temperature gradients in the tropics are determined by elevation (e.g. Forero-Medina et al., 2011), divergence among clades of *Ceroxylon* is evident by elevational segregation. Even if a simple examination of the elevational distributions (Fig. 1E) could suggest this hypothesis, relationships between clades with complementary subsets of the
ancestral climatic space strongly support the hypothesis that divergence is driven by climatic niche segregation, something that has rarely been statistically demonstrated (but see Chaves et al., 2011) and is probably an important aspect of orogeny-driven diversification.

Phylogenetic niche conservatism in Ceroxylon has allowed the colonization of the northern Andes due to its pre-adaptation to cool and moist habitats. Ceroxylon has conserved its ancestral climatic niche by occupying cool habitats, which became increasingly available by Andean uplift during the last 8 Myr. Adaptation to cool habitats is present throughout Ceroxylon and in J. australis, indicating that it is easier to move than to evolve (Donoghue, 2008), as also seen in other cold-adapted plant lineages (e.g. Halenia Borkh.: von Hagen & Kadereit, 2003; Lupinus L.: Hughes & Eastwood, 2006). However, within this pattern of overall biome conservatism, the strong environmental gradient over short distances provided by tropical mountains such as the Andes has allowed Ceroxylon to diversify by climatic niche segregation among clades. In more general terms, it contributes to our understanding of the altitudinal replacement of related species in the enormous plant diversity within the Andean biodiversity hotspot.

CONCLUSIONS AND FUTURE RESEARCH PROSPECTS

Ceroxylon provides an example of Neogene colonization of the northern Andes following Andean uplift facilitated by a combination of phylogenetic niche conservatism and niche segregation within these overall limits. During colonization, clades evolved adaptations to specific climatic conditions, primarily along variables related to temperature (elevation). Thus, the climatic niche was segregated among clades in the phylogenetic tree, supporting a process of elevational segregation. In contrast to niche differentiation at the clade level, the climate niche was broadly conserved among species within a clade. This study provides a clear example of the role played by mountains in the origin of tropical plant communities.

Furthermore, adaptation to cool habitats is also conserved in all Neotropical Ceroxyleae, diverging in the Oligocene. Such tolerance may have been a precondition not only for migration of Ceroxyleae into the high-elevation tropics, but also for dispersal of the tribe across the Southern Hemisphere, perhaps via Antarctica, which did indeed harbour palms during the relevant time window (Pross et al., 2012). Testing for phylogenetic niche conservatism at different levels in phylogenetic trees will provide important insights into ecological trait evolution, diversification, historical migration and assembly of tropical communities, and of their contribution to other species pools.

ACKNOWLEDGEMENTS

This study was funded by the European Research Council under the European Union’s Seventh Framework Programme FP7-PALMS program (FP7-ENB-2007-I) and by the Dirección Nacional de Investigación (DIB) of the Universidad Nacional de Colombia (Project Code 15154). W.D.K. was supported by the Danish Council for Independent Research–Natural Sciences (grant 11-106163) and a University of Amsterdam starting grant. C.D.B. was supported by the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013, ERC Grant Agreement no. 331024). J.C.S. was supported by the European Research Council (ERC-2012-StG-310886-HIST-FUNC). We thank Henrik Balslev, Ricardo Callejas, Yisela Figueroa, Natalia Arcila, and members of the Plant Biodiversity and Adaptation Research Group at Institut de Recherche pour le Développement in Montpellier. We also thank William Baker and anonymous reviewers for their valuable insights into the manuscript. This study is dedicated to the memory of Jean-Christophe Pintaud, who carried out this project together with us and shared our passion for Andean plant biodiversity and biogeography and was our valued friend, colleague and mentor.

REFERENCES


© 2016 The Linnean Society of London, Botanical Journal of the Linnean Society, 2016, 182, 303–317


SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** Time-calibrated Bayesian maximum clade credibility species tree of Ceroxylon of three chloroplast (atpH-atpI, atpH-aptF, ndhf) and one nuclear (PRK) genes.

**Table S1.** Tyukey mean pair comparisons (95% confidence level) of PCA axes values for either temperature or precipitation BioClim variables between pairs of clades; P-values indicating significant differences among pairs of clades for the four PCA axes are shown.

**Table S2.** Tyukey mean pair comparisons (95% confidence level) of altitude of records between pairs of clades; P-values indicating significant differences among pairs of clades for the four PCA axes are shown.