Herbarium-based phylogenomics reveals that the Andes are a biogeographic barrier for *Otoba* (Myristicaceae), an ecologically dominant Neotropical tree genus

Laura Frost[1], Daniel A. Santamaría-Aguilar[1], Daisy Singletary[1], & Laura P. Lagomarsino[1,2]

[1]Shirley C. Tucker Herbarium, Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70808, USA

[2]Author for correspondence (e-mail: llaogmarsino1@lsu.edu)

Abstract

• Premise of the study — Universal probesets for targeted sequence capture have facilitated phylogenomic research into diverse plant groups with limited genomic resources, including from low-quality DNA typical of herbarium specimens. Here, we leverage the Angiosperms353 loci to infer the first phylogeny of *Otoba* (Myristicaceae), a Neotropical tree genus that is ecologically dominant in low-to-mid elevation wet forests, exclusively from herbarium specimens.

• Methods— We use a combination of Angiosperms353 loci, obtained via targeted sequence capture, and plastid sequences to resolve the phylogeny of *Otoba* using concatenated and species tree methods. We subsequently use this phylogeny to infer biogeography and trait evolution using phylogenetic comparative methods.

• Key results— Recovery success of loci is correlated with age of herbarium specimens and average annual precipitation. Despite a large amount of missing data, we resolve the phylogeny of *Otoba* into three major subclades, each structured by geography. We show that *Otoba*'s crown radiation occurred on the western slopes of the Andes in the late Miocene, and from there, migrated into Central America at least twice; the genus was only able to cross to the eastern slopes of the Andes a single time. Trait evolution has been dynamic across vegetative and reproductive traits, with multiple origins of most discrete traits investigated, including ecologically important aril color.

• Conclusions — *Otoba* is recent, rapid radiation whose evolution is tied to landscape change, including Andean uplift, in the northern Neotropics. Its dynamic morphological evolution is consistent with sorting of ancestral traits during recent speciation events. In one of the first herbariomic studies exclusively using herbarium tissue from specimens collected in the wet tropics, this study demonstrates the promise of Angiosperms353 loci in resolving shallow species-level relationships, even from low-quality DNA.

Keywords

Amazonia; Chocó; herbariomics; Magnoliales; museum-based research; natural history collections; Neotropics; phylogeny; seed dispersal

The Northern Neotropics experience a dramatic recent geological history, punctuated by periods of rapid mountain uplift in the Andes (Hoorn et al., 2010) and Central America (De Boer et al., 1995) and a potentially gradual closing of the Isthmus of Panama (Bacon et al., 2015; O'Dea et al., 2016). Among the more notable geographic features of the Neotropics are the Andean mountains of western South America, the longest north-to-south oriented mountain chain in the world and the second tallest of any globally. These mountains began their rise in the Paleocene, but major bursts of mountain building occurred more recently, 4 to 12 My ago. This uplift not only changed local topography, but also impacted continental-scale climate and the entire landscape of the Neotropics (Hoorn et al., 2010). It is thus not surprising that the Andes are known to be important in structuring biogeography and species relationships in plants that occur in montane Andean habitats (Pennington et al., 2010; Särkinen, Pennington, et al., 2012; Lagomarsino et al., 2016; Hoorn et al., 2019), as well as in extra-Andean plants, including lowland tropical rainforest plants of the Amazon basin (Antonelli et al., 2009; Dick et al., 2012).

Both the Isthmus of Panama and the Andean mountains are important geological features of the northern Neotropics that promote speciation in allopatry following long distance dispersal (Gentry, 1982; Antonelli et al., 2009). Given their height and extreme habitat heterogeneity, the Andean mountains are a particularly important barrier to species movement, especially to tropical species that typically have relatively limited environmental preferences (Janzen, 1967). While there are widespread species that occur on either side of the Andes, including *Cordia alliodora* (Boraginaceae) (Rymer et al., 2013), *Symphonia globulifera* (Clusiaceae) (Dick and Heuertz, 2008), and *Schizolobium parahyba* (Fabaceae) (Turchetto-Zolet et al., 2012), it is more common that genera, not species, have trans-Andean distributions. Many of these groups are also found in Central America, which may have been facilitated by the closing of the Isthmus of Panama between 3 and 15 Mya (Coates and Stallard, 2013; Bacon et al., 2015). Myristicaceae, a pantropical family of mid-canopy and canopy trees with high species richness in the Neotropics (ter Steege et al., 2006), includes multiple genera with trans-Andean distributions.

Across their full diversity (comprising 21 genera and *ca.* 500 species), Myristicaceae are notable for their importance in ethnobotany, including as food plants (e.g., nutmeg and mace, *Myristica fragrans*), timber species (e.g., *Virola surinamensis*), and hallucinogenic drugs (e.g., epená, *Virola* sp. (Alrashedy and Molina, 2016)). Myristicaceae also have multidimensional ecological importance. For example, due to their large, arillate seeds, they are some of the most important food sources in the lowland tropics for large-bodied birds, such as toucans and mot-mots, primates, and bats, which, in turn, act as important seed dispersers (Forget et al., 2000;

Russo, 2003; Giraldo et al., 2007; Melo et al., 2009; Moreira et al., 2017). Though pollination is less well studied, the small, usually imperfect flowers, of which thousands can be in bloom at a single time on an individual plant (Kanstrup and Olesen, 2000), are known to be generalist pollinated by various small insects including beetles, flies, and thrips in *Myristica* (Armstrong and Irvine, 1989; Sharma and Armstrong, 2013), *Virola* (Jardim and Mota, 2007). A similar set of small, generalist pollinators is likely common throughout the distribution of the family. Myristicaceae are usually dioecious, with individual plants producing either carpellate or staminate flowers, which results in differential resource allocation, with pistillate plants (which give rise to the fruit) investing more in stem growth than staminate plants (Queenborough et al., 2007a). Further, the six Neotropical genera of Myristicaceae, *Bicuiba*, *Compsoneura*, *Iryanthera*, *Osteophloeum*, *Otoba*, and *Virola*, are important components of wet tropical forests at low to mid-elevations, and individual species can be among the most abundant in lowland tropical rainforests (ter Steege et al., 2006). Neotropical Myristicaceae are an important system for understanding the ecological processes that allow species co-existence in hyperdiverse communities in the western Amazon Basin (Queenborough et al., 2007b; c).

Despite its relatively low species richness, *Otoba* (Fig. 1) has among the broadest ecological tolerances of all Neotropical Myristicaceae. This genus of *ca.* 10 species are distributed from Nicaragua to Brazil, with the highest species richness in low Andean montane forests and lowland rainforests, especially of the Chocó region and western Amazon (Santamaría-Aguilar, Jiménez, et al., 2019). They are abundant in low Andean montane forest, and include the highest elevation occurrence of any member of Myristicaceae (Jaramillo-Vivanco and Balslev, 2020). However, species can also be found in lowland rainforests, and *Otoba* is one of the ten most abundant genera in western Amazonia (ter Steege et al., 2006; Guevara Andino et al., 2017). Individual species can be some of the most common in many forests, including *O. parvifolia* in Madre de Díos, Peru (Pitman et al., 2017; Swamy, 2017) and Madidí, Bolivia (Macía, 2008), *O. glycycarpa* in Yasuní, Ecuador (Guevara Andino et al., 2017), and high várzea forest of the Amazonian floodplain in Brazil and Bolivia (Wittmann et al., 2006). *Otoba parvifolia*, a wide-spread species of the Western Amazon, shows high intraspecific genetic differentiation (Honorio Coronado et al., 2019).

Otoba is distinct among Myristicaceae in many regards. Like other members of the nutmeg family, *Otoba* is characterized by a strong aromatic scent from essential oils, a pagoda-like growth form (i.e., “Myristicaceous growth”, or Massart’s model (Hallé et al., 1978)), dioeciousness with small, trimerous flowers (Armstrong and Tucker, 1986) (Fig. 1A-C), red, dilute latex (Fig. H), and a characteristic valvate capsule that opens to reveal a large, arillate

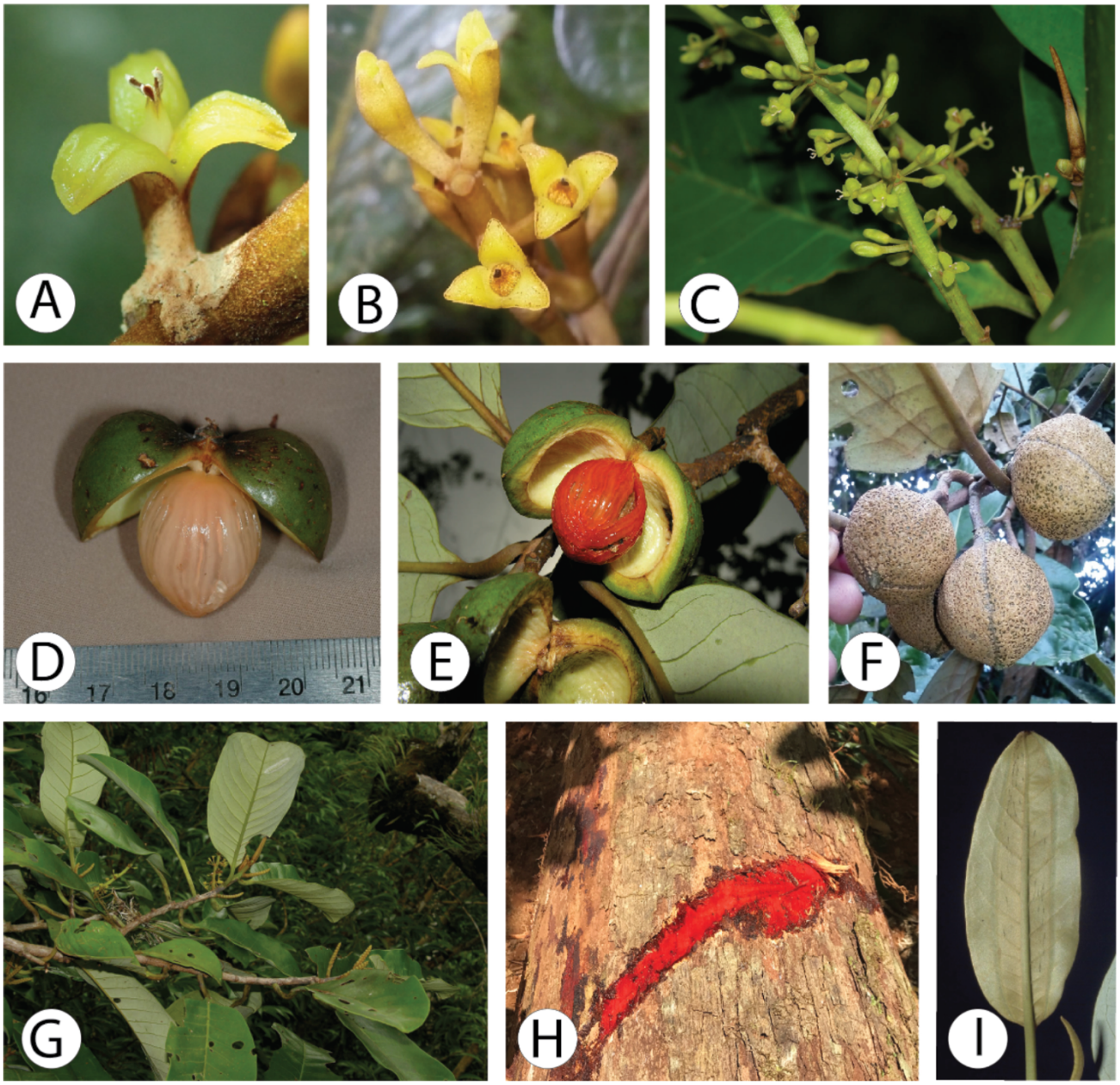

Figure 1. Morphological diversity of *Otoba*. A-C) Floral diversity. A) Staminate and B) pistillate flowers of O. gordoniifolia; C) Infloresence of Central American O. novogranatensis. D-F) Fruit diversity. D) Fruit from South American O. novogranatensis showing whitish aril and E) from Central American O. novogranatensis showing red aril. F) Unopened capsules of O. gordoniifolia. G-I) Vegetative diversity. G) Branch and H) stem cut of Central American O. novogranatensis, the latter showing characteristic red exudate. I) Leaf of O. parvifolia, showing vernation lines. (Photo credits: A, B, and F by Rudy Gelis, downloaded from iNaturalist with permission; C, E, G, and H by Reinaldo Aguilar; D by Timothy Paine; and I by John Janovec.)

112
113

seed (Fig. 1D-E). However, *Otoba* can be distinguished from these genera by a variety of traits. Within Neotropical Myristicaceae, *Otoba* is notable for its low-montane distribution (vs. the more common lowland rainforest) and seeds that most commonly have white arils (vs. typically brightly colored, as in mace). It has bifacial pollen with continuous tectum on the distal pole and reticulate tectum on the proximal pole (Sauquet and Le Thomas, 2003), a unique set of characters within Myristicaceae that is more similar to African members of the family than other Neotropical genera. It is also the only genus of Neotropical Myristicaceae with conduplicate vernation (Fig. 1).

These differences are perhaps not surprising given its relatively distant relationship to other Neotropical Myristicaceae. Myristicaceae is sister to the rest of Magnoliales (Qiu et al., 2006; Soltis et al., 2007; Massoni et al., 2014), a magnoliid clade that represents one of the oldest angiosperm orders (Magallón et al., 2015). Compared to other families within its order, Myristicaceae have low levels of genetic diversity. This may be a product of the relatively recent origin of extant Myristicaceae, compared to other Magnoliales, in the Miocene or late Oligocene (15–18 My: (Doyle et al., 2004; Massoni et al., 2015). This is corroborated by recent molecular dating analyses across angiosperms that suggest that the family originated at least 20 My after the crown-group of Magnoliales at 109 My (Magallón et al., 2015). Despite these young molecular age estimates, fossil evidence suggests that the family has existed since at least the early Eocene, though it is possible that *Myristicacarpum chandlerae* represents a stem lineage (Doyle et al., 2008). Within the family, Myristicaceae is split into three major clades: the Malouchoids, Pycnathoids, and Myristicoids (Doyle et al., 2004). While most Neotropical members of this family are included in the Myristicoids, *Otoba* is nested within the otherwise African Pycnathoids. This, with the relatively young crown age of Myristicaceae, suggests that it represents the product of long-distance dispersal to the Neotropics from Africa, followed by in-situ diversification (Sauquet et al., 2003; Doyle et al., 2004).

Despite their ecological importance and past work into the phylogeny of Magnoliales and Myristicaceae, there is no species-level phylogeny of *Otoba* to date. Further complicating this endeavor, species of *Otoba* are morphologically similar and, as is common for tropical trees, poorly represented in herbaria, making systematic treatment challenging (Gentry, 1979; Bebber et al., 2010).Like many other primarily tropical groups, Myristicaceae has few genomic resources, with only a single publicly available transcriptome (*Myristica fragrans*; (Carpenter et al., 2019). Luckily, the recent development of universal probe sets that target loci of phylogenetic utility across angiosperms, including Angiosperms353, has facilitated the phylogenomic analysis of such understudied lineages. Further, the molecular technique they

rely upon, targeted sequence capture, is robust to the low-quantity, low-quality DNA from herbarium specimens. This is important, as inclusion of herbarium specimens is often necessary to achieve robust taxon sampling for clades that grow in regions that are difficult to reach or to collect in for a variety of reasons, including international collaborations, government permitting, or regional unrest (Rabeler et al., 2019). While herbarium specimens serve as a valuable resource to improve taxonomic sampling, DNA from herbarium specimens collected in the wet tropics have been shown to perform relatively poorly due to preservation issues (Brewer et al., 2019) posing a challenge to robust genetic sampling. We take advantage of the universal nature of the Angiosperm353 probeset to infer the first phylogenetic hypotheses for *Otoba*. This is also the first phylogenomic study that relies exclusive on DNA extracted from herbarium specimens collected in the wet tropics. Using this phylogenetic framework, we discuss implications for biogeography and trait evolution of *Otoba*.

MATERIALS AND METHODS

Taxon Sampling

Twenty accessions of *Otoba* representing nine (*O. acuminata, O. cyclobasis, O. glycycarpa, O. gordoniifolia, O. gracilipes, O. latialata, O. novogranatensis, O. parvifolia, and O. vespertilio*) of the ten accepted species and two undescribed species (*Otoba* sp. nov.) were sampled. All accessions came from herbarium specimens; voucher information may be found in Appendix S1 (see the Supplemental Data with this article). Herbarium acronyms follow Index Herbariorum (Thiers, constantly updated: http://sweetgum.nybg.org/science/ih/). To serve as outgroups, data from the following transcriptomes available on 1KP project (Carpenter et al., 2019; One Thousand Plant Transcriptomes Initiative, 2019); <https://db.cngb.org/onekp/>) were gathered for Myristicaceae (*Myristica fragrans*), the broader Magnoniales (*Magnolia maudiae*, *Annona muricata*), and Laurales (*Cassytha filiformis*, *Sassafras albidum*, and *Persea borbonia*).

DNA extraction, library prep, target enrichment, and sequencing

Dried leaf tissue was weighed to obtain 500 mg, and tissue was homogenized using an MP Biomedicals FastPrep-24TM 5G Instrument. DNA extraction followed a modified sorbitol extraction protocol (Štorchová et al., 2000). Double-stranded DNA concentration was quantified using an Invitrogen Qubit 4 Fluorometer, and fragment size was assessed on a 1% agarose gel. For samples with a high concentration of large fragments (>800 bp), the DNA was sheared using a Bioruptor Pico (Diagenode Inc., Denville, New Jersey, United States) to obtain an average fragment size of ~ 500 bp. Library preparation was carried out using KAPA Hyper Prep

and KAPA HiFi HS Library Amplification kits (F. Hoffmann-La Roche AG, Basel, Switzerland) and with iTru i5 and i7 dual-indexing primers (BadDNA, University of Georgia, Athens, Georgia, United States). Library preparation with KAPA Hyper Prep followed the manufacturer's protocol (KR0961 – v8.20), except for the following modifications: reaction volumes were halved (25 µL starting reaction, instead of 50 µL), and bead-based clean-ups were performed at 3X volume rather than 1X volume to preserve more small fragments from degraded samples. Library amplification reactions were performed at 50 µL. Target enrichment was carried out using the MyBaits Angiosperms353 universal probe set (Däicel Arbor Biosciences, Ann Arbor, MI; (Johnson et al., 2019)). Target enrichment followed the modifications to the manufacturer's protocol outlined in (Hale et al., 2020); i.e., pool of 20-24 samples and RNA baits diluted to ¼ concentration). Unenriched DNA library was added to the cleaned, target enriched pool to increase the amount of off-target, chloroplast fragments in the sequencing library. DNA libraries were sent to Novogene Corportation Inc., (Sacramento, California, United States) for sequencing on an Illumina Hiseq 3000 platform with 150 bp paired-end reads.

Sequence processing, assembly, and alignment

Raw sequence reads were demultiplexed by Novogene Corportation Inc., (Sacramento, California, United States). Adapter sequence removal and read trimming were performed using illumiprocessor v2.0.9 (Faircloth et al., 2012; Faircloth, 2016), a wrapper for trimmomatic v0.39 (Bolger et al., 2014). The default settings were used and reads with a minimum length of 40 bp kept.

HybPiper v. 1.3.1 (Johnson et al., 2016) was used to assemble and extract target regions. Read mapping, contig assembly and coding sequence extraction were performed running the reads_first.py script. The intronerate.py script was run to extract introns and intergenic sequences flanking targeted exons. The retrieve_sequences.py script was run first with the "dna" argument to extract coding regions and subsequently with the "supercontig" argument to extract both coding and non-coding regions as a single concatenated sequence for each target gene. Individual genes were aligned using MAFFT v. 7.310 (Katoh and Standley, 2013). Alignments were visually inspected in AliView v. 1.18.1 (Larsson, 2014) to identify alignment errors, assembly errors, and areas that were difficult to align. Alignment errors were manually corrected and assembly errors, as well as areas that were difficult to align, were removed from individual alignments. Outgroup sequences were added to cleaned alignments and aligned using MUSCLE v.3.8.31 (Edgar, 2004) as the default aligner program in AliView

(Larsson, 2014). Summary statistics on gene alignments were obtained using AMAS (Borowiec, 2016), including length, missing data, and number of parsimony informative sites.

Off-target chloroplast reads were extracted using FastPlast v1.2.6 (https://zenodo.org/record/973887). For all samples there was insufficient data to produce a fully-assembled chloroplast genome. The SPAdes-assembler built into FastPlast iteratively used k-mer lengths of 55, 87, and 121. Assembled contigs from the iteration using k-mer length 87 were mapped to a reference plastome obtained from GenBank (Clark et al., 2016): Horsfieldia pandurifolia (GenBank accession number NC_042225.1). Once mapped, contigs were cleaned by eye to remove assembly errors before generating a consensus sequence. Consensus sequences for each sample were aligned visually against the *Horsfieldia* plastome, as alignment algorithms performed poorly with the large amounts of missing data over long sequences. Plastomes for *Annona muricata* (MT742546.1), *Cassytha filiformis* (MF592986.1), *Magnolia maudiae* (MN990580.1), and an unverified plastome for *Myristica yunnanensis* (MK285565.1) were added as additional outgroups.

Assessing the impact of specimen age and climate on capture success

The collection year of each voucher specimen was recorded and the annual precipitation (mm) at the collection locality extracted from the WorldClim 2.0 30s Bioclimatic variable layer (Fick and Hijmans, 2017) using R package raster (Hijmans et al., 2015). Linear regressions were performed for collection year and annual precipitation number versus number of target loci recovered and average sequence length recovered for each sample to determine if the age of specimen and/or the amount rainfall at the collection locality affected the success of target sequence capture. The relationship between age, precipitation, and the number of ungapped basepairs in cleaned chloroplast sequences was also examined to assess the effect of these factors on off-target sequence capture.

Phylogenetic analyses

Gene tree reconstruction- Maximum likelihood (ML) estimation of gene trees was performed for each nuclear locus, a dataset with all nuclear loci combined, the chloroplast genome, and a dataset with chloroplast and nuclear data combined. Alignments were processed with trimAl (Capella-Gutiérrez and Silla-Martínez, 2009) assigning a gap threshold of 15% or 20% to each column, depending on the number of taxa in the alignment. Thresholds were chosen to maintain columns with data for four or more individuals. Alignments were analyzed using RAxML v8.2.12 (Stamatakis, 2014) under the GTR model with optimization of substitution rates and site-specific

evolutionary rates. For combined datasets and the chloroplast, analyses were first run with all individuals and the program RogueNaRok v.1.0 (Aberer et al., 2011) was used to identify individuals that negatively impacted phylogenetic inference. Individuals identified by RogueNaRok and or those with little data (total bp <1% of aligned length) were excluded from further analyses.

Multispecies Coalescent **-** Trees were generated under the multispecies coalescent model in ASTRAL-III (Zhang et al., 2018). Twenty random bootstrap trees were selected from each gene tree analysis of the Angiosperm353 loci, and used as the input for ASTRAL.

Divergence time estimation

Divergence times were estimated on the ML chloroplast-nuclear combined tree using penalized likelihood via the chronos() function in the R package ape (Paradis and Schliep, 2019). Crown ages from the literature (Magallón et al., 2015; Massoni et al., 2015) for Laurales + Magnoliales, Laurales, Magnoliales, and Myristicaceae were applied as secondary calibrations (Table 2). Because (Massoni et al., 2015) presented five different calibration schemes, and therefore five sets of dates for each node, we calculated the mean age for our calibrations as the average of their mean ages estimated by BEAST across the different schemes. The minimum and maximum for each node were selected as the youngest and oldest date, respectively, in the 95% confidence interval of any scheme across the different analyses. To estimate the error surrounding dates at uncalibrated nodes, like the crown age of *Otoba*, the median/mean, minimum and maximum value from each study was applied as an absolute age at the corresponding nodes.

Ancestral State Reconstruction

Ancestral character estimation was performed using the morphological characters scored in a recently published taxonomic revision *of Otoba* (Jaramillo-Vivanco and Balslev, 2020), including 10 discrete characters and 18 continuous characters. Character states were applied to the ML chloroplast and nuclear combined topology calibrated to maximum ages in (Magallón et al., 2015) and trimmed to include one representative of each species. *Otoba vespertilio* was not included in ancestral state reconstructions due to the uncertainty surrounding its phylogenetic placement.

Because *Otoba* is small genus, and our phylogeny included 8 species, the following discrete characters were simplified from (Jaramillo-Vivanco and Balslev, 2020) to reduce the

number of possible characters states: petiole wingedness; pubescence on the underside of leaves; anther shape and attachment; aril color. The degree to which the petiole is winged was simplified from four categories ("obscurely", "somewhat", "winged", and "not winged"; (Jaramillo-Vivanco and Balslev, 2020) to three (obscurely to somewhat winged, winged, and not winged). Pubescence on the underside of leaves was similarly simplified from "pubescent", "glabrescent", "somewhat pubescent", and "densely pubescent" (Jaramillo-Vivanco and Balslev, 2020) to just two character states: glabrescent to somewhat pubescent and pubescent to densely pubescent; *O. novogranatensis* ranges from glabrescent to densely pubescent and was thus coded as occupying both states. Anther shape and attachment did not vary (i.e, anthers were either globose and dorsally attached or reniform and basally attached, but never reniform and dorsally attached or globose and basally attached), so they were combined into a single character. Aril color was altered to have one state representing pale arils ("white", "white-yellow", and "yellow") and one representing darker arils ("orange-reddish" and "red"). Aril lacination was excluded from this study as information is only available for *O. acuminata*, *O. lehmannii*, and *O. vespertilio*. For continuous characters, the midpoint was taken for measurements given as a range in (Jaramillo-Vivanco and Balslev, 2020) and two-dimensional traits (e.g., ovary size [length x width (mm)] and seed size [length x width (mm)]) were separated into two traits (e.g., ovary length and ovary width).

Since *O. novogranatensis* was recovered as polyphyletic, the state for the broadly described *O. novograntensis* was applied to both *O. novogranatensis* populations in our tree, with the exception of aril color, which we know differs across inferred lineages (see (Santamaría-Aguilar, Jiménez, et al., 2019). Ancestral characters were estimated using the R package phytools (Revell, 2012); the ace() and the fastAnc() functions were used for discrete and continuous traits, respectively.

Biogeographic inference

We modeled biogeographic movements using BioGeoBEARS (Matzke, 2013, 2014; Massana et al., 2015) implemented in RASP v.4.0 (Yu et al., 2015). The same time-calibrated, trimmed topology used in ancestral state reconstructions was used for biogeographic reconstructions. To better understand how major geologic events, like the closure of the Isthmus of Panama and Andean orogeny, correlate with biogeographic events in the evolutionary history of Otoba, movement both between continents, as well as distribution on either side of the Andes were modeled. Each species was coded for occurrence in (A) Central America, (B) South America, or (AB) both. Species were also coded for their distribution on (A) the western side of

Andes, including the Darién gap and Central America or (B) the eastern side of the Andes and western Amazonia. Six biogeographic models were tested with BioGeoBEARS (Matzke, 2013, 2014; Massana et al., 2015); the DIVA-like model was selected for reconstruction of continental movements and the BAYArea-like model with jump dispersal was selected for reconstruction of distribution around the Andes. A maximum of two ancestral areas was allowed for both analyses.

RESULTS

Summary statistics of data assembly— The number of Angiosperm353 loci captured, average sequence length, number of ungapped basepairs of chloroplast DNA (cpDNA) for each sample, collection year, and annual precipitation at the collection locality are listed in Table 1. A heatmap of the percent of the reference protein length recovered for each sample at each locus can be found in Appendix S2 and summary statistics for each locus in Appendix S3.

Capture success from herbarium specimens—We found variable success from hybrid enriched target sequence capture across samples. Half of the 20 samples submitted for sequencing recovered fewer than 10 Angiosperm353 loci; only 3 samples recovered more than 100 loci (Table 1). Success in gathering off-target chloroplast data did not necessarily correspond to success in capturing nuclear loci. For example, the sample with the most nuclear data– *O. novogranatensis*_WS36336 with 217 of the 353 targeted loci– did not recover useful chloroplast data. On the other hand, nearly half of the chloroplast genome was obtained for *O. parvifolia*_MS1182, despite recovering only 6 nuclear loci.

Specimens used for DNA extraction were collected between 1983 and 2015, and annual precipitation ranged from 1449 mm/year to 3870 mm/year (Table 1). There was a positive correlation with collection year and both the number of loci recovered and the average sequence length recovered--more recently collected specimens tended to recover longer sequences for more loci--though only the correlation between age and average sequence length was significant (Fig. 2). *Otoba novogranatensis*_LG20482 was an outlier; despite being the oldest specimen, this sample performed well in both target sequence capture and off-target capture of the chloroplast (Table 1; Figure 2). Annual precipitation was negatively, but not significantly, correlated with the number of loci, average sequence length. Higher rainfall in the collection locality reduced the performance of extracted DNA with target sequence capture. There was very little correlation between age or rainfall and off-target capture success (Fig. 2); the outlying sample *Otoba novogranatensis*_LG20482 did influence results. Without this

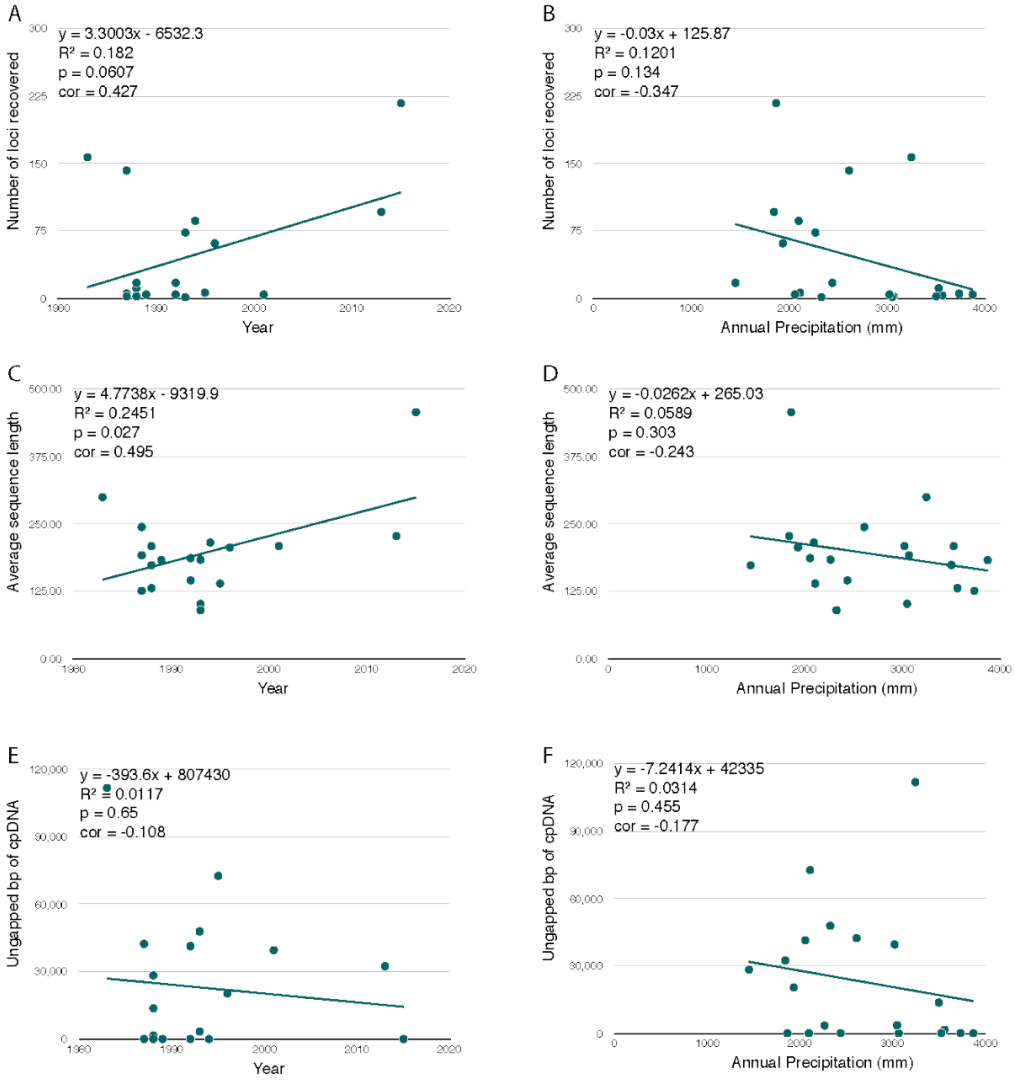

Figure 2. Linear regressions for (A) collection year and (B) annual precipitation versus number of loci recovered; (C) collection year and (D) annual precipitation versus average sequence length; (E) collection year and (F) annual precipitation versus number of ungapped basepairs (bp) of chloroplast DNA (cpDNA) recovered from the 20 samples of *Otoba*. The equation, R squared value, and p value (Pr(>|t|) for the linear regression as well as correlation of the two variables are shown in the top left corner of each graph.

sample, the weak negative correlation between collection year and off-target capture success became weakly positive and remained insignificant (cor=0.136, p-0.578). Meanwhile, the weak negative correlation between annual precipitation and off-target capture success became more negative, but remained insignificant (cor=-0.431, p=0.065).

Due to large amounts of missing data in both nuclear and chloroplast regions, the following samples were excluded from all analyses: *O. gracilipes*_DC884, *O. novogranatensis*_EB500, *O. parvifolia*_DN9151, *O. sp. nov.*_RC5752, *O. sp. nov.*_JP16902.

Otoba vespertilio, a recently described species from the Caribbean coast of Costa Rica and Panama (Santamaría-Aguilar, Jiménez, et al., 2019), was included in some analyses, despite large amounts of missing data, in an effort to place the species in the phylogeny (Figs. 3 and 4; Appenidces S4-S6). Maximum likelihood analyses were run with and without this species (Figs. 3 and 4; Appendices S4-S6). Overall, we were able to include 8 of the 10 described species of *Otoba* in phylogenetic analyses and 7 in comparative analyses.

Phylogenetic analyses

We found strong support for the monophyly of *Otoba* within Myristicaceae, albeit with limited outgroup sampling within the family (nuclear ML: 100, nuclear MSC: 1, chloroplast: 78, cp and nuclear combined ML: 100). Though the exact relationship between species varies across our individual analyses, we have strong support for three major clades of *Otoba* (Figs. 3 and 4). The first includes *O. acuminata*, *O. cylclobasis, O. gordoniifolia*, *O. latialata*, and *O. novogranatensis* specimens collected in Central America (nuclear ML: 94, nuclear MSC: 0.9, chloroplast: n/a, cp and nuclear combined ML: 100). The second includes *O. parvifolia* and *O. glycycarpa* (nuclear ML: n/a, nuclear MSC: 1, chloroplast: n/a, cp and nuclear combined ML: 75). The third includes individuals of *O. novogranatensis* collected in South America (nuclear ML: n/a, nuclear MSC: n/a, chloroplast: 63, cp and nuclear combined ML: 88).

Maximum likelihood analyses of concatenated nuclear loci recovered a poorly-supported grade of *O. glycycarpa* and *O. parvifolia* successively sister to a well-supported clade including *O. cyclobasis*, *O. latialata*, *O. gordoniifolia*, and *O. novogranatensis* (Fig. 3A). With *O. vespertilio* included, the widespread species *O. parvifolia* is paraphyletic with respect to *O. vespertilio*. Samples for another widespread species, *O. novogranatensis*, are monophyletic, but the nuclear dataset only included individuals from Central America. The ASTRAL-III tree also found a clade with *O. cyclobasis*, *O. latialata*,*O. gordoniifolia*, and *O. novogranatensis* with high support as well as a strongly-supported clade of *O. glycycarpa*, *O. parvifolia*, and *O. vespertilio* (Fig. 1B). *Otoba parvifolia* remains non-monophyletic in the ASTRAL-III results; however, relationships within the clade are poorly supported (Fig. 3B).

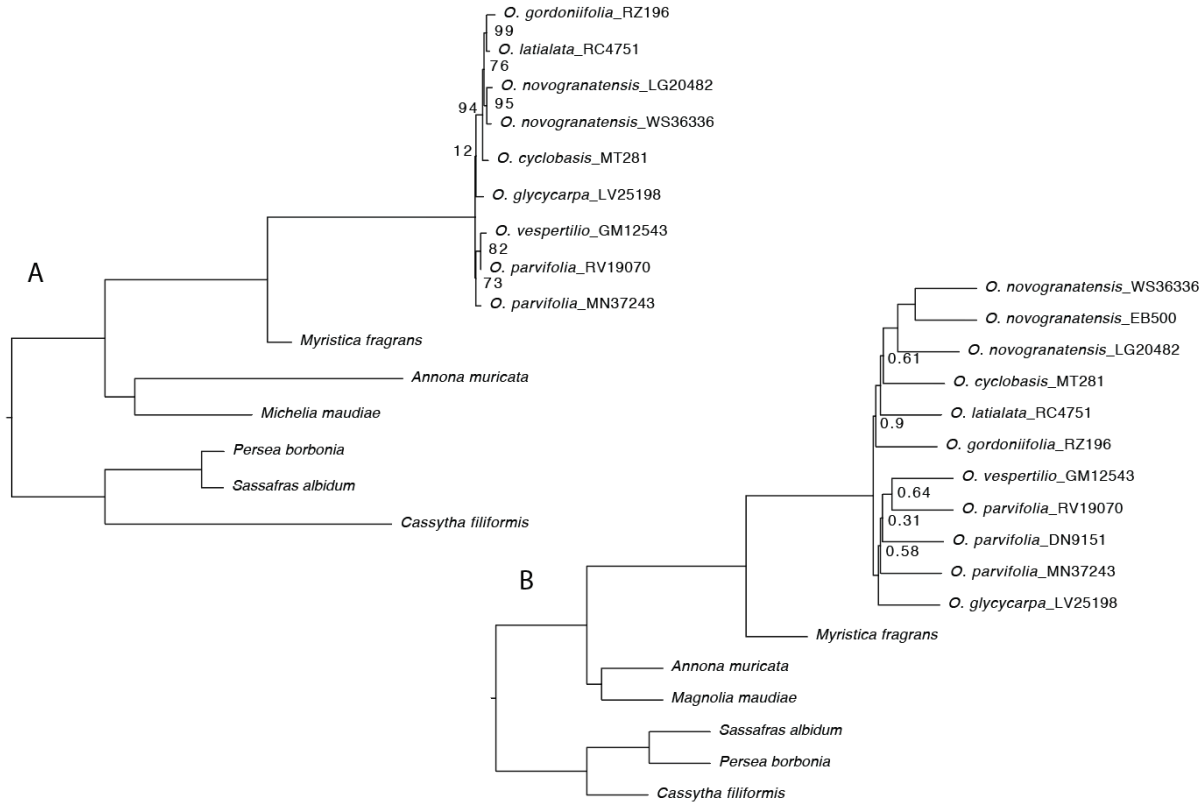

Figure 3. Results of phylogenetic analyses with nuclear data; (A) ML analysis of concatenated loci using RAxML and (B) multispecies coalescent analysis of 20 random bootstrap trees from each gene tree analysis using ASTRAL. Support values are listed at nodes for those with <100 bootstrap support and <1.0 local posterior probability for the RAxML and the ASTRAL tree, respectively.

The topology of the ML chloroplast tree did not agree with either of the nuclear topologies; however, relationships in the chloroplast tree are, overall, poorly supported (Appendix S5). The strongly-supported nodes in the ingroup included *Otoba* as a clade (bootstrap support=78) and the sister pair *O. acuminata* and *O. novogranatensis* (CA, sample LG20482; bootstrap support=85). Albeit with weak support along the backbone, the South American samples of *O. novogranatensis* are inferred to be sister to the rest of the genus. The clade of South American *O. novogranatensis* samples is resolved as sister to the rest of the genus in the combined chloroplast and nuclear analyses (bootstrap support=88; Fig. 4). The topology for the remainder of *Otoba* from the chloroplast and nuclear combined dataset is congruent with analyses of nuclear data (Figs. 3 and 4). The clade including *O. cyclobasis*, *O. latialata*, *O. gordoniifolia*, and the Central American *O. novogranatensis* is again found with high support; a well-supported clade including *O. glycycarpa*, *O. parvifolia* is also recovered (Fig. 4).

However, when *O. vespertilio* was included in the combined dataset, *O. vespertilio* nests within the South American *O. novogranatensis* rather than *O. parvifolia* (Appendix S5).

Otoba lehmannii was not sampled in this study and we did not recover sufficient data to include *O. gracilipes* in our phylogenetic analyses. Based on the geographical structure of clades within *Otoba*, these species likely either belong to the larger western Andean/Central American clade or in the South American *O. novogranatensis* clade.

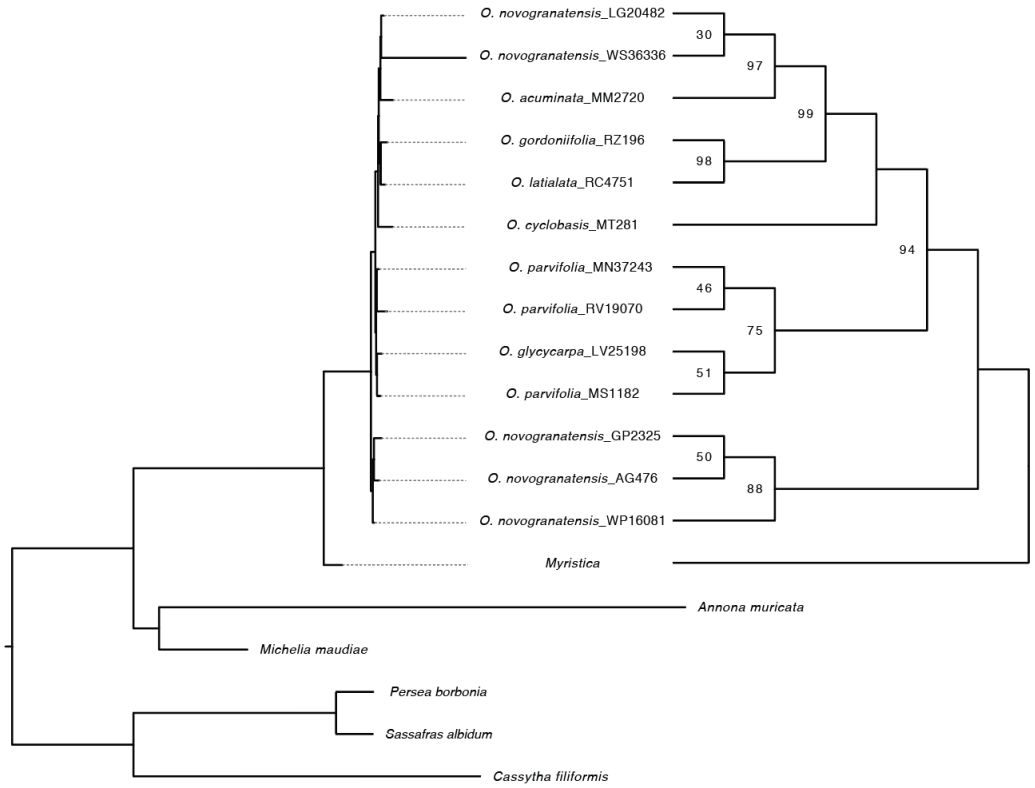

Figure 4. Results of ML analyses of concatenated chloroplast and nuclear data. The tree on the right shows branch lengths; the cladogram on the left shows the branching pattern for the ingroup + *Mryistica* and support values at nodes with <100 bootstrap support (all outgroup relationships were fully supported).

Divergence Time Estimation

Divergence times estimated from different calibrations across studies (Magallón et al., 2015; Massoni et al., 2015) were largely congruent (Table 2; Appendices S7-S12). The mean estimated ages based on (Massoni et al., 2015) were older than those based on (Magallón et al., 2015) and had a broader estimated minimum and maximum range. This is expected as (Massoni et al., 2015) estimated ages under five different timelines for the crown radiation of

angiosperms: 130 Ma, 140 Ma, 150 Ma, 170 Ma, and 200 Ma. The estimated age of angiosperms from (Magallón et al., 2015) was 139.4 Ma. All estimates support the radiation of *Otoba* in the late Miocene (Fig. 6). The crown age for *Otoba* based on calibrations from Magallón et al. (2015) is estimated to be 7.28 Ma (7.04-7.51 Ma); whereas dates based on (Massoni et al., 2015) are 8.69 Ma (6.5 -11.36 Ma). The divergence between the western Andean/Central American clade and the eastern Andean/Amazonian clade is inferred to have occurred around 6.54 Ma (6.32-6.75 Ma) and 7.67 Ma (5.98-9.83 Ma) based on (Magallón et al., 2015) and (Massoni et al., 2015), respectively.

Ancestral State Reconstruction

Many of the discrete characters exhibited frequent transitions between states along the tree (Figure 5; Appendices S13-S22). These frequent shifts in character state among closely-related individuals, combined with short divergence times along the backbone of the tree, produced near-equivocal reconstructions of character states at ancestral nodes (Appendices S9-S18). Similarly, closely-related species exhibited marked differences in continuous traits (Figure 5; Appendices S23-S40). All estimations of continuous characters resulted in a value at the root close to the average of the observed states in the genus (Appendices S19 - S36).

Traits with a resolved evolutionary history include: the presence of secondary to tertiary intramarginal veins, the shape and fusion of filaments, the presence of extrastaminal discs, the presence of prickles at the apex of the seed, and aril color (Appendices S10, S11, S12, S17, and S18). Within the more-resolved character histories, multiple derivations of traits are inferred. The presence of secondary to tertiary intramarginal veins emerged at least three times; they are present in *O. gordoniifolia*, populations of *O. novogranatenis*, and *O. lehmannii* (not sampled). Fused, bottle-shaped filament columns are the most common state in *Otoba*. Filament shape transitioned from the tapered bottle shape to cylindrical in *O. cyclobasis*. Fusion of filaments was lost multiple times in populations of *O. novogranatensis*. Extrastaminal discs evolved twice: once in *O. acuminata* and once in *O. cyclobasis*. Prickles at the apex of the seed were partially lost in *O. cyclobasis* and *O. parvifolia* and lost completely in *O. glycycarpa*; *Otoba vespertilio* (not included in ancestral reconstructions) also displays a reduction in seed prickles. Finally, arils are ancestrally pale (white to yellow) in *Otoba*; red arils have evolved at least once in Central American *O. novogranatensis*. One additional species not included in ancestral reconstructions, *O. gracilipes*, also has red arils.

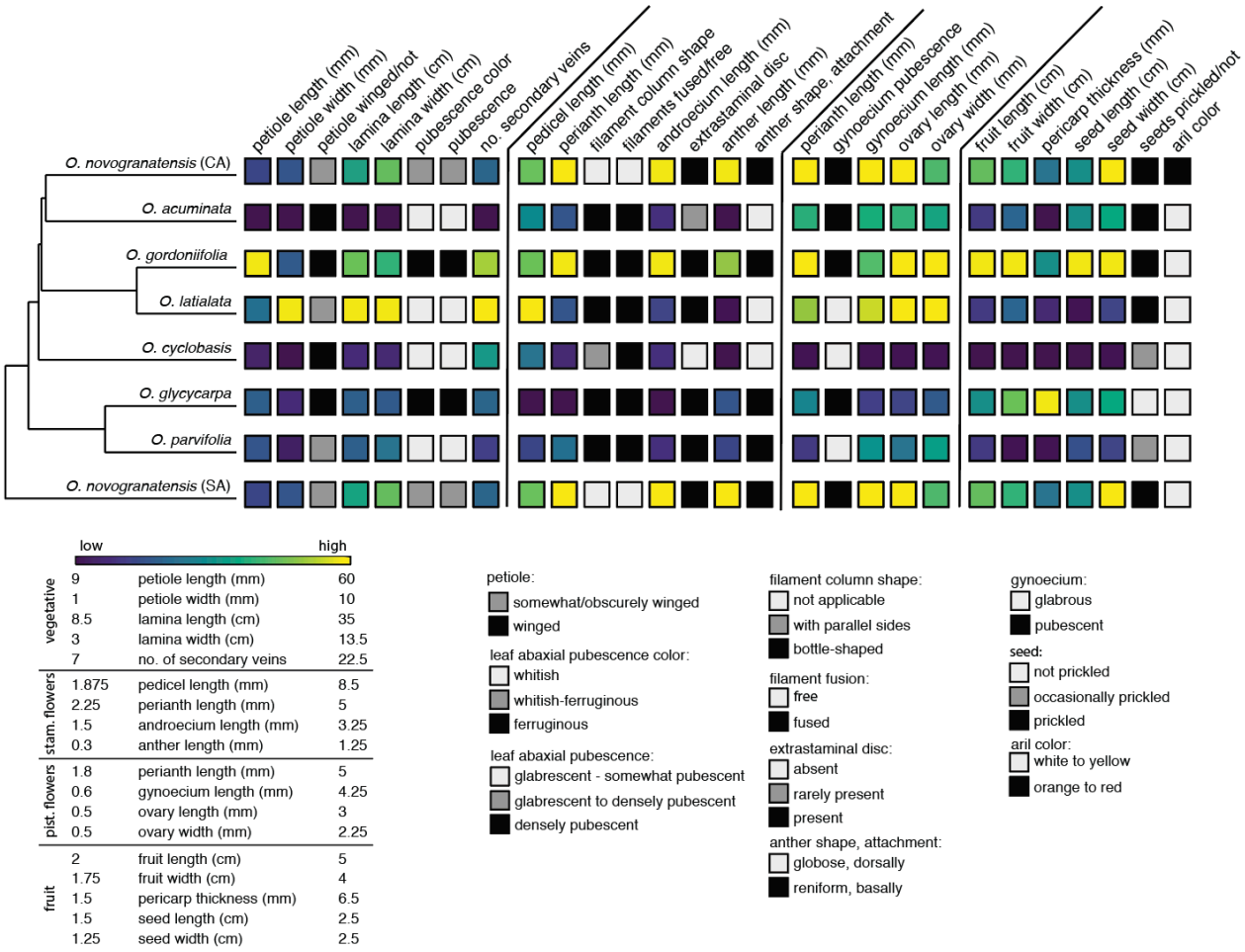

Figure 5. Summary of tip states from ancestral character reconstruction for species included in the phylogeny. Traits are sectioned into vegetative, staminate floral, pistillate floral, and fruit traits. Discrete traits are color-coded in grayscale with the scoring system for each trait described below. Continuous characters are coded along a colored gradient from low values (dark purple) to high values (yellow). The numerical range for each trait is listed below; color values for each tip were extracted from the output of ancestral character estimations in R (Appendices S9 – S37).

Biogeographic Reconstruction

South America and, more specifically, the western side of the Andes is inferred to be the ancestral area of *Otoba* (Fig. 6). Jump dispersal to the eastern side and the Amazonian basin best explains the divergence of the *O. glycycarpa*-*O. parvifolia clade* (Fig. 6B). In the western-Andean clade including *O. acuminata* and *O. cyclobasis*, a South American ancestor is inferred with expansion into Central America via widespread ancestors (Fig. 6A).

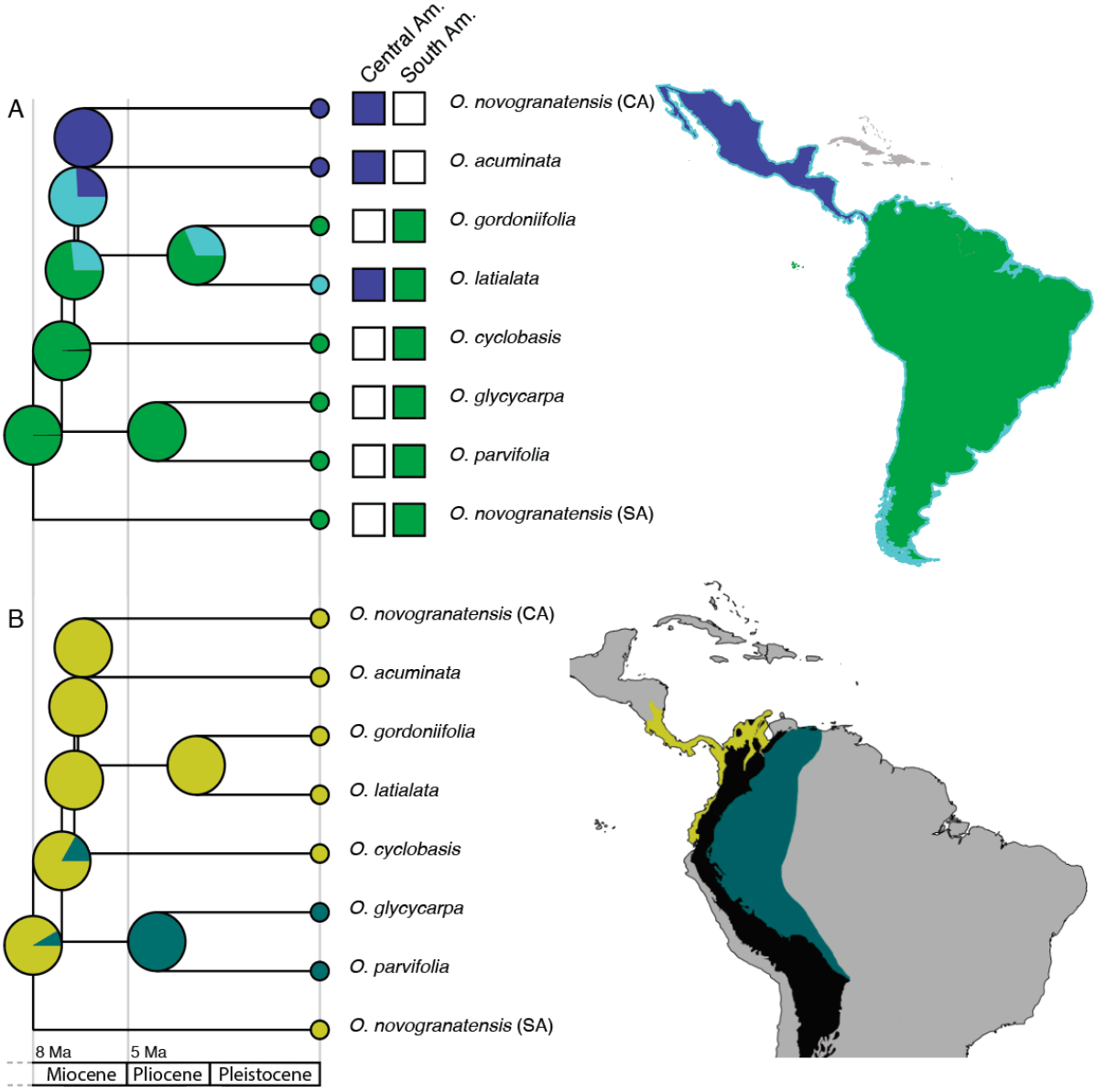

Figure 6. Ancestral area reconstructions for (A) Central America versus South America and (B) western Andes and Central America versus eastern Andes and the Amazon. Pie charts at nodes display the probability that the ancestor occupied a given range. Maps to the left of each tree are color coded to correspond with the geographic areas coded on the tree. In (A), green = South America, royal blue = Central America, and light blue = both Central and South America; the distribution of each species in Central and/or South America is additionally reflected in the color-coded boxes at the tips of the tree. In (B), chartreuse = western Andes and Central America, and teal = eastern Andes and the Amazon; the Andes are represented in black.

463

464

DISCUSSION

There were many challenges, both methodological and biological, inherent to our goal of inferring the first phylogeny of the ecologically dominant Neotropical tree genus *Otoba*. Our use of a universal probe set meant lower specificity in target capture, and that variation within loci would potentially be low. Methodological challenges extended to our exclusive use of degraded, low quantity DNA extracted from herbarium specimens— a challenge exacerbated by *Otoba*'s distribution in low- to-mid-elevation moist tropical forests, where specimens are particularly difficult to preserve and dry. Biologically, *Otoba*'s Andean-centered distribution suggests that it is likely a young clade with a phylogeny characterized by short branch lengths and young speciation events, a result borne out in our analyses. Despite these significant challenges, we were able to produce a framework phylogeny for *Otoba*— the first phylogeny for any Neotropical group of Myristicaceae. Our phylogenetic results consistently included two subclades: 1) an eastern Andean/Amazonian clade comprising *O. parvifolia and O. glycycarpa* and 2) a northern Andean/Central American clade comprising *O. acuminata*, *O. cyclobasis*, *O. gordoniifolia*, *O. latialata*, and some individuals of *O. novogranatensis*. The 6-9 Ma divergence of the eastern Andean/Amazionian clade from the western Andean/Central American clade suggests that long-distance dispersal across the Andes, rather than vicariance via mountain uplift, produced the present distribution.

Landscape evolution and dispersal limitation drive biogeography in Otoba

The Neotropics have a complex geological history (Hoorn et al., 2010), including the rapid uplift of the Andes and the closure of the Isthmus of Panama. These events have had dramatic impacts on the evolution of the taxa that inhabit this region (Hughes et al., 2006; Antonelli et al., 2009; Bacon et al., 2015). We find compelling evidence that both have played important roles in the shaping biogeographic history of *Otoba*.

Otoba is a product of long-distance dispersal from Africa to the Neotropics. It is likely that the dispersal event that resulted in stem lineage *Otoba* was from western Africa, where the closest relatives of *Otoba*, *Pycnanthus and Ceolocaryon*, occur today, to eastern South America (Doyle et al., 2004) in the last 19-7.2MA (Table 1). During this time period, range expansion across the northern Neotropics would have been facilitated by a more-or-less contiguous swathe of lowland rainforest and a relatively low height of the Andes, including a low elevation gap in the region that comprises the current border area between Ecuador and Peru (Hoorn et al., 2010). By the time that crown *Otoba* originated *ca.* 7.2Ma, we infer that this lineage was

restricted to the western side of the Andes; the long, empty stem lineage of *Otoba* is consistent with high rates of past extinction explaining this pattern. Our ancestral state reconstructions and divergence time estimates point to a western Andean origin of *Otoba* (Fig. 6) in the late Miocene to early Pliocene. This is a region of origin shared with relatively few other Neotropical tree groups, including a clade of Annonaceae (Pirie et al., 2018); more tree clades that have been investigated have Amazonian origins, including the *Brownea* clade (Fabaceae; (Schley et al., 2018) and Neotropical Chrysobalanaceae (Chave et al., 2020). Following establishment, *Otoba* seem to be one of few lineages that experience *in situ* diversification within the Chocó biogeographic region (Pérez-Escobar et al., 2019).

The Andes structure the ranges of species and clades in crown *Otoba*. Species either occur on the eastern slopes of Andes, extending into the western Amazon basin, or on the western slopes of the Andes and/or in Central America. The same pattern is observed within subclades, with closest relatives sharing distribution on one side of the Andes. During the time that elapsed between the dispersal from Africa and the origin of crown *Otoba,* the northern Andes gained approximately half their elevation (Garzione et al. 2017) and analogs to modern montane cloud forests formed (Hughes, 2016; Martínez et al., 2020). Thus, even though they had not yet reached their full height (Hoorn et al., 2010), the Andes would have represented a significant barrier to dispersal for low- to-mid-elevation groups like *Otoba*. Consistent with this scenario, we infer only a single dispersal event from the western side of the Andes to the eastern side; this dispersal occurred in the late Miocene to early Pliocene, resulting in two widespread, lowland species, *O. glycycarpa* and *O. parvifolia* (Fig. 6). It is unlikely that Andean uplift served as a vicariant event that split widespread populations; instead, we posit that long-distance dispersal across the mountains, likely mediated by bird or mammal seed dispersers, facilitated this disjunct distribution.

Movement between Central and South America has occurred more frequently than movement across the Andes during the evolutionary history of *Otoba.* Range expansion and long-distance dispersal have both played a role in movement into Central America. A widespread distribution in South and Central America is inferred for the most recent common ancestor of *O. acuminata*, *O. novogranatensis* (CA), *O. gordoniifolia*, and *O. latialata* in the late Miocene (Fig. 6). The current, widespread distribution of *O. latialata* in the Chocó-Darién moist forest from Colombia to Panama suggests the common ancestor of the clade may have occupied a similar distribution. If this is the case, a pattern of subset sympatry—when one daughter lineage inherits the ancestral range and the other daughter(s) inherit a portion of the ancestral range (Ree et al., 2005)—is consistent with our inferred biogeographic history. An

additional long-distance dispersal event is assumed for *O. vespertilio*. While the phylogenetic position of *O. vespertilio* differs across our analyses (Fig. 3; Appendix S6), all possible placements are consistent with it representing a long distance dispersal event from South America to Central America. These events occurred within the last 10 million years, a time-frame that supports the role of the closure of the Isthmus of Panama in facilitating these movements (Montes et al., 2012; Bacon et al., 2013).

The observed biogeography patterns are likely explained by dispersal limitation. *Otoba*'s relatively large seeds are dispersed birds, primates, and bats (Giraldo et al., 2007; Nuñez-Iturri and Howe, 2007; Melo et al., 2009; Santamaría-Aguilar, Jiménez, et al., 2019), with a potential role of small mammals as secondary seed dispersers (Forget et al., 2002). These groups can often be dispersal limited, including across riverine barriers and fragmented habitats (Eberhard and Bermingham, 2005; Ripperger et al., 2013; Boubli et al., 2015) (though this is not always the case; see (Holbrook, 2011). The potential for dispersal limitation is observed within communities in which the relatively large seeds of *Otoba parvifolia* are dispersed at low frequency over typically short distances (Terborgh et al., 2011). On a continental scale and over evolutionary time, this has resulted in relatively few long distance biogeographic movements. *Otoba*'s two migrations into Central American were potentially facilitated by the continuous land bridge of more-or-less suitable habitat, while cold high-elevation habitats likely prevented more frequent traversing of the Andes. Further, the range of *Otoba* does not occupy all of the suitable habitat it presumably could based on distributions of extant species. Otoba does not occur in other low to mid-elevation moist forest habitats in the Neotropics, like the Atlantic coast forest in Brazil and the Caribbean islands, a pattern mirroring ecologically similar and closely related *Virola* (Santamaría-Aguilar, Aguilar, et al., 2019). Again, dispersal limitation may explain *Otoba*'s absence from these regions: the relatively large seeds of *Otoba* likely make dispersal events over water barriers or large stretches of unsuitable terrestrial habitat uncommon compared to groups that are dispersed by wind (Pérez-Escobar et al., 2017) or migratory passerine birds (Nathan et al., 2008). This is likely coupled to time limitation, given the young age of the genus and the stochastic nature of long-distance dispersal events (Nathan et al., 2008). However, high levels of seed-set, both in closed canopy forests and in treefall gaps, may make *Otoba* an effective colonizer of new habitats once they do arrive in a new region (Myster, 2020).

Dynamic evolution of morphological traits is common

We observe a dynamic pattern of character evolution across a broad suite of traits in *Otoba*. While frequent transitions between morphological characters often results in unresolved

ancestral state reconstructions, mapping tip states reveals that no two species share a set of discrete traits (Fig. 5). Further, even though ancestral states were equivocal, the majority of the discrete traits that we investigated underwent convergent evolution within the genus, including: winged petioles, pubescence presence and color, filament fusion, anther shape and attachment, and gynoecium pubescence (Fig. 4). Coupled with short divergence times between speciation events after the crown age of *Otoba* (Fig. 6; Appendices S7 and S8), the observed variation across all traits is consistent with rapid morphological evolution coinciding with rapid lineage divergence upon establishment in South America. This pattern may also be explained by the sorting of ancestral variation (Pease et al., 2016), and it is likely that hemiplasy or parallel evolution underlie morphological evolution. Ecological opportunities on a new continent may have served to differentially select standing ancestral variation, resulting in the repeated evolution of many traits in *Otoba*.

An unconfirmed, but likely case of convergent evolution is aril color. Most species in Myristicaceae produce red arils that cover seeds to varying degrees. These bright, red arils that contrast with the green and brown of the capsule and seeds (Fig. 1E) serve as an attractant to frugivores who consume the nutritious aril, either with or without consuming the seed (Howe and Vande Kerckhove, 1981; Gautier-Hion et al., 1985). Most species of *Otoba*, however, produce white to yellow arils (Fig 1F); these pale colors still contrast the seed and capsule and attract frugivores (Gautier-Hion et al., 1985; Wheelwright and Janson, 1985). Two species have red arils: *O. gracilipes* and the Central American members of *O. novagranatensis* (supported as a distinct lineage in our analyses). Our field observations suggest that these differences in aril color coincide with differences in texture: red arils are generally thick and waxy, while whitish arils are thin and gelatinous. On a broad phylogenetic scale, this represents convergent evolution of red arils within Myristicaceae and likely serves as a shared ecological signal with other Neotropical members of the family, including the widespread *Virola*. Within the genus, bright red arils have presumably evolved twice independently (Fig. 5), though the exact phylogenetic placement of these independent origins remains unknown as *O. gracilipes* was not included in our analyses. We also see convergence in seed size and shape (Fig. 5), additional traits that are likely related to the mechanics of seed dispersal.

Convergence in fruit traits, including seed size and aril color and texture, may be linked to a specialized dispersal syndromes in *Otoba*. Differences in overall seed morphology throughout the genus, including size, aril color and texture, potentially represent specialization with different classes of dispersers. For example, species with bright red arils, including *Virola surinamensis*, are consumed by birds (Howe and Kerckhove, 1980; Howe and Vande

Kerckhove, 1981), while species with whitish arils are consumed by bats (Melo et al., 2009; Santamaría-Aguilar, Jiménez, et al., 2019); while more comparative field studies are needed to confirm the most effective dispersers across species, these observations are broadly consistent with global patterns of mammal- and bird-dispersed fruits (Sinnott-Armstrong et al., 2018). While seed dispersal is not directly tied to reproductive isolation, theoretical models support mutualisms between animal dispersers and flowering plants coupled with repeated habitat fragmentation as a mechanism that promotes speciation during range contractions, allowing for coexistence of diverged populations upon habitat reunification (Kakishima et al., 2015). This is a likely potential scenario in *Otoba*, given that major mountain-building events were occurring the northern Andes and, along with them, reorganization of river systems, were occurring during the late Miocene and early Pliocene, when many of the divergence times between species are inferred (Hoorn et al., 1995, 2010; Struth et al., 2015). Fluctuating landscapes and fluvial barriers at the time may have fragmented populations, promoting divergence between those populations and securing species boundaries upon secondary contact. Supporting this, large rivers are important barriers to dispersal and gene flow in tropical birds (Hayes and Sewlal, 2004; Burney and Brumfield, 2009; Fernandes et al., 2014; Oliveira et al., 2017; Sandoval-H et al., 2017); by extension, they should also be significant barriers to the plants that they disperse as well (Nazareno et al., 2017; Dambros et al., 2020).

The need for future phylogeography of widespread species

Otoba includes a combination of narrowly endemic and widespread species. We find strong evidence for polyphyly of one widespread taxon, *O. novogranatensis,* and for paraphyly of another, *O. parvifolia*. This non-monophyly is likely both a product of biological processes (e.g., very large population sizes maintained over very large distances in *O. novogranatensis* (Pennington and Lavin, 2016) and an artifact of insufficient taxonomic study (Lagomarsino and Frost, 2020).

The polyphyletic *Otoba novogranatensis* is one of the most collected species of its genus. A preliminary revision of the resulting specimens shows variation across many traits. For example, South American specimens differ from Central American specimens in their thicker pericarp, pubescent ovaries, anthers that can be unfused to the base (de Candolle, 1856; Jaramillo et al., 2004; Jaramillo-Vivanco and Balslev, 2020), and generally white arils. Thus, the two distinct lineages of *O. novogranatensis* that we resolve (Fig. 4), corresponding to South American and Central American accessions, are supported by morphology. There are additional differences within Central American *O. novogranatensis* that are likely to be taxonomically

635 informative as well, including the size, shape, and pubescence of leaves, as well as the length
636 of staminate inflorescences, perianth, and anthers varies across the Pacific (e.g., *R. Aguilar 638*
637 [INB/CR]) and Caribbean slopes (e.g., *I. Chacón* 1360 [INB/CR]) of Costa Rica; montane
638 populations within Costa Rica exhibit additional differentiation. Together, these observations
639 suggest that phylogeographic analysis of *O. novogranatensis* is necessary to adequately
640 determine the number of lineages that are currently described under this umbrella taxon. This
641 would facilitate future taxonomic efforts to recircumscribe this species complex to reflect
642 evolutionary relationships, which will likely entail the description of at least one new Central
643 American species.
644 Similarly, the paraphyly of *O. parvifolia* is not surprising, but for different reasons. Our
645 phylogenetic results suggest that *O. glycycarpa* has recently diverged from *O. parvifolia*, and
646 has since become widespread as well. Our taxon sampling includes multiple accessions of *O.*
647 *parvifolia* and a single accession of *O. glycycarpa*; these species determinations were based on
648 differences in pericarp thickness, gynoecium pubescence, and color of foliar pubescence
649 (Jaramillo et al., 2000). In all analyses, these accessions are each others' closest relatives,
650 though *O. parvifolia* is almost always resolved as a grade (Fig. 3, 4; but see Appendix S4 where
651 it is resolved as a clade). The geographic structure of these species are unresolved across our
652 analyses; while Bolivian and Peruvian accessions of *O. parvifolia* are often resolved as sister
653 lineages that are more distantly related from Ecuadoran and Brazilian accessions *O. parvifolia*
654 and *O. glycycarpa*, there is variation across our analyses and support for these relationships is
655 often low. A previous study based on limited genetic data has suggested that *O. parvifolia* and
656 *O. glycycarpa* are genetically indistinguishable, with the implication that they represent a single
657 species (Honorio Coronado et al., 2019). This low genetic variability between species, which is
658 consistent with the short branch lengths in our analyses, may be the product of species
659 misidentification, lack of informative variation in the loci used, introgression, or sorting of
660 ancestral variation in the short time since speciation (*ca.* 3.7 Ma). The latter two mechanisms
661 would not be surprising: our analyses suggest that these species are recently diverged,
662 overlapping in both geographic occurrence, and have overall similar morphology. While our
663 sampling did not allow us to test the monophyly of *O. glycycarpa*, morphological evidence
664 supports the distinctness of these species (Jaramillo et al., 2000). Of particular note, *O.*
665 *glycycarpa* has one of the thickest pericarps of all *Otoba*, which has implications for potential
666 efficacy of seed dispersers. We consider *O. parvifolia* to be a distinct species form *O.*
667 *glycycarpa*, and its paraphyly is indicative of its complex evolutionary history (Freudenstein et
668 al., 2016). Future ecological and phylogeographic research could target what allows

morphologically similar closest-related species to co-occur throughout large swathes of western Amazonia.

Targeted sequence capture promotes herbariomics— with caveats

Hybrid-enriched target sequence capture with Angiosperms353 has proven useful for generating phylogenetically informative data at multiple taxonomic scales and from different sources of DNA (i.e., silica-dried tissue versus herbarium specimens; (Brewer et al., 2019; Shee et al., 2020; Valderrama et al., 2020). To our knowledge, our study represents a new milestone for Angiosperms353 phylogenomics: the first exclusively herbariomic dataset for a wet tropical genus. DNA from herbarium specimens collected in the wet tropics performs poorly compared extractions from other climates and silica-gel dried tissue (Brewer et al., 2019). This is because high humidity in moist tropical forests and often remote localities extend drying times and/or delay access to drying apparatuses. Resultantly, collections are often treated with ethanol to prevent mold and decay until they can be dried. Both storage of tissues in humid conditions after collection (Adams, 2011) and preservation in ethanol (Doyle and Dickson, 1987; Pyle and Adams, 1989) degrades DNA, resulting in damaged and fragmented genetic material (Särkinen, Staats, et al., 2012). Even though current short-read sequencing techniques perform well with herbarium specimens as compared to conventional Sanger sequencing (Bakker et al., 2015), the level of degradation common among wet tropical specimens led (Brewer et al., 2019) to recommend the use of silica tissue for wet tropical groups. While we agree that this is a best-practice, as in other herbarium-based studies (Brewer et al., 2019; Shee et al., 2020), we were able to extract useful phylogenomic data from herbarium specimens using Angiosperms353.

Herbarium samples are an increasingly useful source of DNA for phylogenomic studies, extending the utility of our historic natural history collections (Lendemer et al., 2020), but they are not a panacea. Even high efficiency target capture will fail when DNA is low quality, as is typical in herbarium specimens collected in the wet tropics (Brewer et al., 2019), especially when they are collected in ethanol (Särkinen, Staats, et al., 2012)— a common scenario for *Otoba*, and likely universal in the herbarium specimens that we sampled. The specimens we included in our analyses span a breadth of age and environmental conditions at collecting localities. The oldest specimen that we included was 37 years old, collected in 1983; surprisingly, it resulted in the most genetic data (Fig. 2). However, not all of our samples generated useful sequences; we had to remove some of our samples completely due to poor quality reads and apparent contamination following visual inspection of alignments. In other

cases, we were able to extract a handful of useful loci, but at a much lower quality than the majority of our included species.

In agreement with past studies, we found that capture success, as measured by the total number of loci obtained as well as their average length, decreased with increasing age of the specimen; we further found a correlation between environmental conditions (i.e., annual precipitation) at the site of collection and the quality of phylogenomic data (Fig. 2). However, we were only able to capture more than 10 loci from half of our initial 20 samples, and only three *Otoba* samples were successful in recovering over 100 loci. This limited capture success is most likely a product of low input DNA quality and not due to sequence divergence from the probeset, especially considering that we had similarly variable success in capturing the high-copy, off-target chloroplast genome. Plastomes from wet tropical specimens have also been found to have higher fragmentation rates and lower sequencing success as compared to plastomes from specimens collected in other climates (Bakker et al., 2015; Brewer et al., 2019). This compounded with a lower depth of sequencing coverage in the off-target chloroplast regions and resulted in higher sequencing error, which led to conflicting phylogenetic signal and low support. The chloroplast data were able to complement our Angiosperms353 data and allowed the inclusion of additional samples in phylogenetic analyses, but stronger, more consistent signal was recovered in Angiosperm353 target loci.

To extend the potential utility of voucher specimens collected in the future for genomic research, efforts should be made to collect leaf tissue in silica gel or other preservation technique, following best practices, including unique IDs to connect the herbarium voucher to this secondary product (Funk et al., 2017). However, many studies of tropical groups rely on museum specimens, as it is not feasible to collect living material for the taxonomic and/or geographic breadth of many clades. As this is often the case, we urge botanical collectors to include information about how voucher specimens were treated (e.g., specifying that specimens were collected and dried in ethanol). We have demonstrated that even though it may not be ideal, it is still possible to generate phylogenomic datasets that can resolve rapid radiations of wet tropical plant clades using the Angiosperms353 probeset.

Utility of Angiosperms353 as universal loci for plant phylogenomics, even in rapid radiations

Despite the fact that the Angiosperms353 loci were designed to be phylogenetically useful across all angiosperms, we were able to resolve relationships within a clade estimated to be between 6.5–11.3 My old (Table 2). This adds to a growing number of rapid radiations whose

phylogenies have been resolved using these loci (Larridon et al., 2019). This performance goes against a preconception that universal loci are not useful for species-level phylogenomic analysis, probably borne of experience from the low-variation plastid loci that had universal utility via PCR-based Sanger sequencing (Shaw et al., 2005), as well as related calls to move away from the lofty goal of a universal set of loci for all flowering plants and instead develop lineage-specific loci with more phylogenetic utility (Hughes et al., 2006). However, targeted sequence capture allows the isolation of loci without finicky PCR probes at a percent-divergence that allows for efficient capture over substantial evolutionary distances. Further, universal probesets, including Angiosperms353, are designed such that multiple probe sequences for each locus are included to account for sequencing variability across taxa. Instead of being limiting, the universal nature of Angiosperms353 loci are proving to be informative for both very deep (e.g, (Dodsworth et al., 2019) and very shallow (e.g., within populations of rice, (Van Andel et al., 2019) phylogenetic splits. A further attractive aspect of Angiosperms353 data is that it is cost-effective to generate, especially if pre-sequencing molecular labwork is completed by the researcher and not outsourced (Hale et al., 2020).

While we were able to infer relationships in *Otoba* using Angiosperms353 loci, there was very limited variation across our dataset (i.e., the proportion of variable sites in concatenated target loci was 0.285). It has been shown in other taxa that custom bait kits are more informative than Angiosperms353 loci for species-level phylogenetics (Jantzen et al., 2020), though this is not always the case (Larridon et al., 2019). It is thus possible that a custom probe kit designed for *Otoba* and close relatives would have outperformed the Angiosperms353 loci, either alone or in combination. However, developing such custom loci is predicated on the existence of genomic resources, either pre-existing or newly generated, which would have been difficult for *Otoba*. While the number of angiosperms genomic resources is constantly growing (One Thousand Plant Transcriptomes Initiative, 2019), there are still no transcriptomes or nuclear genomes available for *Otoba* and data is limited for Myristicaceae overall: there is a single transcriptome available in the 1KP database (nutmeg, *Myristica fragrans*) and no other genomic resources. Further, we currently do not have access to fresh or silica-dried tissue from *Otoba*. Because this scenario is common— especially in tropical plant groups, which tend to be understudied (Goodwin et al., 2015), the universal utility and subsequent promise of assembling a standardized set of loci across analyses is a very desirable property of the Angiosperms353 loci. To facilitate standardized loci that can be combined across studies, the plant systematics community should strive for the development of searchable, long-term repositories for Angiosperms353 datasets. Any such future repository would further benefit from being linked to

herbarium vouchers, which would help with reproducibility of research as well as continue to extend the use of herbarium specimens in science (Lendemer et al., 2020).

Conclusions

Otoba is a recent, rapid radiation with a trans-Andean distribution whose biogeographic history suggests that key aspects of the Neotropical landscape— namely the Andean mountains and the Isthmus of Panama— have been important barriers to dispersal. It is likely that traversing these barriers has happened only rarely in the history of the genus due to dispersal limitation, a product of its relatively large seeds that rely on large-bodied vertebrates for dispersal. Very short branch lengths separating species coincide with dynamic trait evolution— a pattern consistent with sorting of ancestral variation during rapid lineage diversification. Future research into *Otoba* would ideally tackle phylogeography of widespread species, including two species that we identify as non-monophyletic: *O. novogranatensis* and *O. parvifolia*, both of which can reach very high local abundance in the communities in which they occur. These insights were gained from a relatively small phylogenomic dataset obtained via targeted sequence capture of Angiosperms353 loci from DNA extracted from herbarium specimens. We observed an impact specimen age and the environmental conditions of their collection locality on data quality. While our study highlights the promise of Angiosperms353 in herbarium-based phylogenomics, it demonstrates the challenges inherent in studying rapid radiations broadly, and, more specifically, when using DNA extracted from herbarium specimens collected in the humid tropics.

Acknowledgements

This research was funded by a Louisiana Board of Regents Research Competitiveness Subprogram grant and by the LSU Department of Biological Sciences. We would like to thank the Missouri Botanical Garden (MO) for their access to their important collections, and Reinaldo Aguilar, Rudy Gelis, Timothy Paine, and John Janovec for permission to use their photographs. We thank Brant Faircloth, Matthew Johnson, Carl Oliveros, and Jessie Salter for their guidance in library preparation and Brant Faircloth for access to laboratory equipment. This manuscript benefited from feedback from Laymon Ball, Janet Mansaray, and Diego Paredes-Burneo.

Author Contributions

LAF, DASA, and LPL conceived of the research; LAF and DS collected data and performed analyses; LAF and LPL wrote the manuscript with significant feedback from DASA and DS.

804

805 **Data Availability Statement**

806 Illumina reads will be submitted to the NCBI Sequence Read Archive (SRA) and all other
807 data formats (tree files, alignments, character matrices, etc.) will be uploaded on the Dryad
808 Digital Repository and made available upon publication.

809

810 **Additional Supporting Information may be found online in the supporting information.**
811

812

Literature Cited

Aberer, A. J., D. Krompaß, and A. Stamatakis. 2011. RogueNaRok: An efficient and exact algorithm for rogue taxon identification. *Heidelberg Institute for Theoretical Studies: Exelixis-RRDR-2011--10*.

Adams, R. P. 2011. DNA from herbarium specimens: II. Correlation of DNA degradation with humidity. *Phytologia* 93: 351–359.

Alrashedy, N. A., and J. Molina. 2016. The ethnobotany of psychoactive plant use: a phylogenetic perspective. *PeerJ* 4: e2546.

Antonelli, A., J. A. A. Nylander, C. Persson, and I. Sanmartín. 2009. Tracing the impact of the Andean uplift on Neotropical plant evolution. *Proceedings of the National Academy of Sciences of the United States of America* 106: 9749–9754.

Armstrong, J. E., and A. K. Irvine. 1989. Floral biology of *Myristica insipida* (Myristicaceae), a distinctive beetle pollination syndrome. *American Journal of Botany* 76: 86–94.

Armstrong, J. E., and S. C. Tucker. 1986. Floral development in *Myristica*. *American Journal of Botany* 73: 1131–1143.

Bacon, C. D., D. Silvestro, C. Jaramillo, B. T. Smith, P. Chakrabarty, and A. Antonelli. 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proceedings of the National Academy of Sciences of the United States of America* 112: 6110–6115.

Bakker, F. T., D. Lei, J. Yu, S. Mohammadin, Z. Wei, S. van de Kerke, B. Gravendeel, et al. 2015. Herbarium genomics: plastome sequence assembly from a range of herbarium specimens using an Iterative Organelle Genome Assembly pipeline. *Biological Journal of the Linnean Society* 117: 33–43.

Bebber, D. P., M. A. Carine, J. R. I. Wood, A. H. Wortley, D. J. Harris, G. T. Prance, G. Davidse, et al. 2010. Herbaria are a major frontier for species discovery. *Proceedings of the National Academy of Sciences of the United States of America* 107: 22169–22171.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Borowiec, M. L. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ* 4: e1660.

Boubli, J. P., C. Ribas, J. W. Lynch Alfaro, M. E. Alfaro, M. N. F. da Silva, G. M. Pinho, and I. P. Farias. 2015. Spatial and temporal patterns of diversification on the Amazon: A test of the riverine hypothesis for all diurnal primates of Rio Negro and Rio Branco in Brazil. *Molecular Phylogenetics and Evolution* 82 Pt B: 400–412.

Brewer, G. E., J. J. Clarkson, O. Maurin, A. R. Zuntini, V. Barber, S. Bellot, N. Biggs, et al. 2019. Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of angiosperms. *Frontiers in Plant Science* 10: 1102.

851 Burney, C. W., and R. T. Brumfield. 2009. Ecology predicts levels of genetic differentiation in
852 neotropical birds. *The American Naturalist* 174: 358–368.

853 de Candolle, A. 1856. Ordo CLXIII. Myristicaceae. *Prodromus Systematis Naturalis Regni*
854 *Vegetabilis* 14: 189–208.

855 Capella-Gutiérrez, S., and J. M. Silla-Martínez. 2009. trimAl: a tool for automated alignment
856 trimming in large-scale phylogenetic analyses. *Bioinformatics* 25: 1972–1973.

857 Carpenter, E. J., N. Matasci, S. Ayyampalayam, S. Wu, J. Sun, J. Yu, F. R. Jimenez Vieira, et
858 al. 2019. Access to RNA-sequencing data from 1,173 plant species: The 1000 Plant
859 transcriptomes initiative (1KP). *GigaScience* 8.

860 Chave, J., C. Sothers, A. Iribar, U. Suescun, M. W. Chase, and G. T. Prance. 2020. Rapid
861 diversification rates in Amazonian Chrysobalanaceae inferred from plastid genome
862 phylogenetics. *Botanical Journal of the Linnean Society*: boaa052.

863 Clark, K., I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and E. W. Sayers. 2016. GenBank. *Nucleic*
864 *Acids Research* 44: D67–72.

865 Coates, A. G., and R. F. Stallard. 2013. How old is the Isthmus of Panama? *Bulletin of marine*
866 *science* 89: 801–813.

867 Dambros, C., G. Zuquim, G. M. Moulatlet, F. R. C. Costa, H. Tuomisto, C. C. Ribas, R.
868 Azevedo, et al. 2020. The role of environmental filtering, geographic distance and dispersal
869 barriers in shaping the turnover of plant and animal species in Amazonia. *Biodiversity and*
870 *Conservation*: https://doi.org/10.1007/s10531–020–02040–3.

871 De Boer, J., M. Drummond, M. Bordelon, M. Defant, H. Bellon, and R. Maury. 1995. Cenozoic
872 magmatic phases of the Costa Rican island arc (Cordillera de Talamanca). Geologic and
873 tectonic development of the Caribbean Plate boundary in southern Central America,
874 Geological Society of America, Boulder, CO.

875 Dick, C. W., and M. Heuertz. 2008. The complex biogeographic history of a widespread tropical
876 tree species. *Evolution* 62: 2760–2774.

877 Dick, C. W., S. L. Lewis, M. Maslin, and E. Bermingham. 2012. Neogene origins and implied
878 warmth tolerance of Amazon tree species. *Ecology and Evolution* 3: 162–169.

879 Dodsworth, S., L. Pokorny, M. G. Johnson, J. T. Kim, O. Maurin, N. J. Wickett, F. Forest, and
880 W. J. Baker. 2019. Hyb-Seq for flowering plant systematics. *Trends in Plant Science* 24:
881 887–891.

882 Doyle, J. A., S. R. Manchester, and H. Sauquet. 2008. A seed related to Myristicaceae in the
883 Early Eocene of southern England. *Systematic Botany* 33: 636–646.

884 Doyle, J. A., H. Sauquet, T. Scharaschkin, and A. Le Thomas. 2004. Phylogeny, molecular and
885 fossil dating, and biogeographic history of Annonaceae and Myristicaceae (Magnoliales).
886 *International Journal of Plant Sciences* 165: S55–S67.

887 Doyle, J. J., and E. E. Dickson. 1987. Preservation of plant samples for DNA restriction
888 endonuclease analysis. *Taxon* 36: 715–722.

889 Eberhard, J. R., and E. Bermingham. 2005. Phylogeny and comparative biogeography of
890 *Pionopsitta* parrots and *Pteroglossus* toucans. *Molecular Phylogenetics and Evolution* 36:
891 288–304.

892 Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high
893 throughput. *Nucleic Acids Research* 32: 1792–1797.

894 Faircloth, B. C. 2016. PHYLUCE is a software package for the analysis of conserved genomic
895 loci. *Bioinformatics* 32: 786–788.

896 Faircloth, B. C., J. E. McCormack, N. G. Crawford, M. G. Harvey, R. T. Brumfield, and T. C.
897 Glenn. 2012. Ultraconserved elements anchor thousands of genetic markers spanning
898 multiple evolutionary timescales. *Systematic biology* 61: 717–726.

899 Fernandes, A. M., M. Cohn-Haft, T. Hrbek, and I. P. Farias. 2014. Rivers acting as barriers for
900 bird dispersal in the Amazon. *Revista brasileira de ornitologia* 22: 363–373.

901 Fick, S. E., and R. J. Hijmans. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces
902 for global land areas. *International Journal of Climatology* 37: 4302–4315.

903 Forget, P. M., D. S. Hammond, and T. Milleron. 2002. Seasonality of Fruiting and Food
904 Hoarding by Rodents in Neotropical Forests: Consequences for Seed Dispersal and
905 Seedling Recruitment. *In* D. J. Levey, W. R. Silva, and M. Galetti [eds.], Seed dispersal and
906 frugivory: ecology, evolution and conservation., 241–256. CAB International, Wallingford,
907 UK.

908 Forget, P.-M., T. Milleron, F. Feer, O. Henry, and G. Dubost. 2000. Effects of dispersal pattern
909 and mammalian herbivores on seedling recruitment for *Virola michelii* (Myristicaceae) in
910 French Guiana. *Biotropica* 32: 452–462.

911 Freudenstein, J. V., M. B. Broe, R. A. Folk, and B. T. Sinn. 2016. Biodiversity and the species
912 concept—Lineages are not enough. *Systematic Biology* 66: 644–656.

913 Funk, V. A., M. Gostel, A. Devine, C. L. Kelloff, K. Wurdack, C. Tuccinardi, A. Radosavljevic, et
914 al. 2017. Guidelines for collecting vouchers and tissues intended for genomic work
915 (Smithsonian Institution): Botany Best Practices. *Biodiversity Data Journal*: e11625.

916 Gautier-Hion, A., J.-M. Duplantier, R. Quris, F. Feer, C. Sourd, J.-P. Decoux, G. Dubost, et al.
917 1985. Fruit characters as a basis of fruit choice and seed dispersal in a tropical forest
918 vertebrate community. *Oecologia* 65: 324–337.

919 Gentry, A. H. 1982. Neotropical floristic diversity: Phytogeographical connections between
920 Central and South America, Pleistocene climatic fluctuations, or an acccident of the Andean
921 orogeny? *Annals of the Missouri Botanical Garden* 69: 557.

922 Gentry, A. H. 1979. Transfer of the species of *Dialyanthera* to *Otoba*. *Taxon* 28: 417–417.

923 Giraldo, P., C. Gómez-Posada, J. Martínez, and G. Kattan. 2007. Resource use and seed
924 dispersal by red howler monkeys (*Alouatta seniculus*) in a Colombian Andean forest.
925 *Neotropical Primates* 14: 55–64.

926 Goodwin, Z. A., D. J. Harris, D. Filer, J. R. I. Wood, and R. W. Scotland. 2015. Widespread
927 mistaken identity in tropical plant collections. *Current Biology* 25: R1066–7.

928 Guevara Andino, J. E., N. C. A. Pitman, H. Ter Steege, H. Mogollón, C. Ceron, W. Palacios, N.
929 Oleas, and P. V. A. Fine. 2017. Incorporating phylogenetic information for the definition of
930 floristic districts in hyperdiverse Amazon forests: Implications for conservation. *Ecology and*
931 *Evolution* 7: 9639–9650.

932 Hale, H., E. M. Gardner, J. Viruel, L. Pokorny, and M. G. Johnson. 2020. Strategies for reducing
933 per-sample costs in target capture sequencing for phylogenomics and population genomics
934 in plants. *Applications in plant sciences* 8: e11337.

935 Hallé, F., P. B. Tomlinson, and M. H. Zimmermann. 1978. Architectural variation at the specific
936 level in tropical trees. *Tropical trees as living systems*: 209–221.

937 Hayes, F. E., and J.-A. N. Sewlal. 2004. The Amazon River as a dispersal barrier to passerine
938 birds: effects of river width, habitat and taxonomy. *Journal of Biogeography* 31: 1809–1818.

939 Hijmans, R. J. 2020. raster: Geographic Data Analysis and Modeling. R package version 3.0-
940 12. https://CRAN.R-project.org/package=raster

941 Holbrook, K. M. 2011. Home range and movement patterns of toucans: Implications for seed
942 dispersal: Toucan movement patterns. *Biotropica* 43: 357–364.

943 Honorio Coronado, E. N., K. G. Dexter, M. L. Hart, O. L. Phillips, and R. T. Pennington. 2019.
944 Comparative phylogeography of five widespread tree species: Insights into the history of
945 western Amazonia. *Ecology and Evolution* 9: 7333–7345.

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

948 Hoorn, C., R. van der Ham, F. de la Parra, S. Salamanca, H. ter Steege, H. Banks, W. Star, et
949 al. 2019. Going north and south: The biogeographic history of two Malvaceae in the wake
950 of Neogene Andean uplift and connectivity between the Americas. *Review of palaeobotany*
951 *and palynology* 264: 90–109.

952 Hoorn, C., F. P. Wesselingh, H. ter Steege, M. A. Bermudez, A. Mora, J. Sevink, I. Sanmartín,
953 et al. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution,
954 and biodiversity. *Science* 330: 927–931.

955 Howe, H. F., and G. A. Kerckhove. 1980. Nutmeg dispersal by tropical birds. *Science* 210: 925–
956 927.

957 Howe, H. F., and G. A. Vande Kerckhove. 1981. Removal of wild nutmeg (*Virola surinamensis*)
958 crops by birds. *Ecology* 62: 1093–1106.

959 Hughes, C. E. 2016. The tropical Andean plant diversity powerhouse. *The New Phytologist* 210:
960 1152–1154.

961 Hughes, C. E., R. J. Eastwood, and C. Donovan Bailey. 2006. From famine to feast? Selecting
962 nuclear DNA sequence loci for plant species-level phylogeny reconstruction. *Philosophical*
963 *Transactions of the Royal Society B* 361: 211–225.

964 Jantzen, J. R., P. Amarasinghe, R. A. Folk, M. Reginato, F. A. Michelangeli, D. E. Soltis, N.
965 Cellinese, and P. S. Soltis. 2020. A two-tier bioinformatic pipeline to develop probes for
966 target capture of nuclear loci with applications in Melastomataceae. *Applications in plant*

sciences: e11345.

Janzen, D. H. 1967. Why mountain passes are higher in the tropics. *The American Naturalist* 101: 233–249.

Jaramillo, T. S., P. Muriel, and H. Balslev. 2004. Myristicaceae. *In* G. W. Harling, and L. Andersson [eds.], Flora of Ecuador, 1–101.

Jaramillo, T. S., P. Muriel, W. A. Rodrigues, and H. Balslev. 2000. Myristicaceae novelties from Ecuador. *Nordic journal of botany* 20: 443–447.

Jaramillo-Vivanco, T. S., and H. Balslev. 2020. Revision of *Otoba* (Myristicaceae). *Phytotaxa* 441: 143–175.

Jardim, M. A. G., and C. G. da Mota. 2007. Biologia floral de *Virola surinamensis* (Rol.) Warb. (Myristicaceae). *Revista Árvore* 31: 1155–1162.

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

Johnson, M. G., L. Pokorny, S. Dodsworth, L. R. Botigué, R. S. Cowan, A. Devault, W. L. Eiserhardt, et al. 2019. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-Medoids clustering. *Systematic Biology* 68: 594–606.

Kanstrup, J., and J. M. Olesen. 2000. Plant-flower visitor interactions in a neotropical rain forest canopy: community structure and generalisation level. *In* Totland, Ø., Armbruster, W.S., Fenster, C., Molau, U., Nilsson, L.A., Olesen, J.M., Ollerton, J., Philipp, M., Ågren, J. [ed.], The Scandinavian Association for Pollination ecology honours Knut Fægri, 33–42. The Norwegian Academy of Science and Letters 39, Oslo.

Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.

Lagomarsino, L. P., F. L. Condamine, A. Antonelli, A. Mulch, and C. C. Davis. 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *The New Phytologist*.

Lagomarsino, L. P., and L. A. Frost. 2020. The central role of taxonomy in the study of Neotropical biodiversity. *Annals of the Missouri Botanical Garden* 105: 405–421.

Larridon, I., T. Villaverde, A. R. Zuntini, L. Pokorny, G. E. Brewer, N. Epitawalage, I. Fairlie, et al. 2019. Tackling rapid radiations with targeted sequencing. *Frontiers in Plant Science* 10: 1655.

Larsson, A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30: 3276–3278.

Lendemer, J., B. Thiers, A. K. Monfils, J. Zaspel, E. R. Ellwood, A. Bentley, K. LeVan, et al. 2020. The Extended Specimen Network: A strategy to enhance US biodiversity collections, promote research and education. *Bioscience* 70: 23–30.

1006 Macía, M. J. 2008. Woody plants diversity, floristic composition and land use history in the
1007 Amazonian rain forests of Madidi National Park, Bolivia. *Biodiversity and Conservation* 17:
1008 2671–2690.

1009 Magallón, S., S. Gómez-Acevedo, L. L. Sánchez-Reyes, and T. Hernández-Hernández. 2015. A
1010 metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity.
1011 *The New Phytologist* 207: 437–453.

1012 Martínez, C., C. Jaramillo, A. Correa-Metrío, W. Crepet, J. E. Moreno, A. Aliaga, F. Moreno, et
1013 al. 2020. Neogene precipitation, vegetation, and elevation history of the Central Andean
1014 Plateau. *Science Advances* 6: eaaz4724.

1015 Massana, K. A., J. M. Beaulieu, N. J. Matzke, and B. C. O'Meara. 2015. Non-null effects of the
1016 null range in biogeographic models: Exploring parameter estimation in the DEC model.
1017 *bioRxiv*: 026914.

1018 Massoni, J., T. L. P. Couvreur, and H. Sauquet. 2015. Five major shifts of diversification through
1019 the long evolutionary history of Magnoliidae (angiosperms). *BMC Evolutionary Biology* 15:
1020 49.

1021 Massoni, J., F. Forest, and H. Sauquet. 2014. Increased sampling of both genes and taxa
1022 improves resolution of phylogenetic relationships within Magnoliidae, a large and early-
1023 diverging clade of angiosperms. *Molecular Phylogenetics and Evolution* 70: 84–93.

1024 Matzke, N. J. 2014. Model selection in historical biogeography reveals that founder-event
1025 speciation is a crucial process in island clades. *Systematic Biology* 63: 951–970.

1026 Matzke, N. J. 2013. Probabilistic historical biogeography: new models for founder-event
1027 speciation, imperfect detection, and fossils allow improved accuracy and model-testing.
1028 *Frontiers of Biogeography* 5.

1029 Matzke, N. J. 2013b. BioGeoBEARS: BioGeography with Bayesian (and Likelihood)
1030 Evolutionary Analysis in R scripts. R package, version 0.2.1. http://CRAN.R-
1031 project.org/package=BioGeoBEARS

1032 Melo, F. P. L., B. Rodriguez-Herrera, R. L. Chazdon, R. A. Medellin, and G. G. Ceballos. 2009.
1033 Small tent-roosting bats promote dispersal of large-seeded plants in a Neotropical forest.
1034 *Biotropica* 41: 737–743.

1035 Moreira, J. I., P. Riba-Hernández, and J. A. Lobo. 2017. Toucans (*Ramphastos ambiguus*)
1036 facilitate resilience against seed dispersal limitation to a large-seeded tree (*Virola*
1037 *surinamensis*) in a human-modified landscape. *Biotropica* 49: 502–510.

1038 Myster, R. W. 2020. Disturbance and response in the Andean cloud forest: a conceptual review.
1039 *The Botanical Review* 86: 119–135.

1040 Nathan, R., F. M. Schurr, O. Spiegel, O. Steinitz, A. Trakhtenbrot, and A. Tsoar. 2008.
1041 Mechanisms of long-distance seed dispersal. *Trends in Ecology & Evolution* 23: 638–647.

1042 Nazareno, A. G., C. W. Dick, and L. G. Lohmann. 2017. Wide but not impermeable: Testing the
1043 riverine barrier hypothesis for an Amazonian plant species. *Molecular Ecology* 26: 3636–
1044 3648.

1045 Nuñez-Iturri, G., and H. F. Howe. 2007. Bushmeat and the fate of trees with seeds dispersed by
1046 large primates in a lowland rain forest in Western Amazonia. *Biotropica* 39: 348–354.

1047 O'Dea, A., H. A. Lessios, A. G. Coates, R. I. Eytan, S. A. Restrepo-Moreno, A. L. Cione, L. S.
1048 Collins, et al. 2016. Formation of the Isthmus of Panama. *Science Advances* 2: e1600883.

1049 Oliveira, U., M. F. Vasconcelos, and A. J. Santos. 2017. Biogeography of Amazon birds: rivers
1050 limit species composition, but not areas of endemism. *Scientific Reports* 7: 2992.

1051 One Thousand Plant Transcriptomes Initiative. 2019. One thousand plant transcriptomes and
1052 the phylogenomics of green plants. *Nature* 574: 679–685.

1053 Paradis, E., and K. Schliep. 2019. ape 5.0: an environment for modern phylogenetics and
1054 evolutionary analyses in R. *Bioinformatics* 35: 526–528.

1055 Pease, J. B., D. C. Haak, M. W. Hahn, and L. C. Moyle. 2016. Phylogenomics reveals three
1056 sources of adaptive variation during a rapid radiation. *PLoS Biology* 14: e1002379.

1057 Pennington, R. T., and M. Lavin. 2016. The contrasting nature of woody plant species in
1058 different Neotropical forest biomes reflects differences in ecological stability. *The New*
1059 *Phytologist* 210: 25–37.

1060 Pennington, R. T., M. Lavin, T. Särkinen, G. P. Lewis, B. B. Klitgaard, and C. E. Hughes. 2010.
1061 Contrasting plant diversification histories within the Andean biodiversity hotspot.
1062 *Proceedings of the National Academy of Sciences of the United States of America* 107:
1063 13783–13787.

1064 Pérez-Escobar, O. A., M. Gottschling, G. Chomicki, F. L. Condamine, B. B. Klitgård, E.
1065 Pansarin, and G. Gerlach. 2017. Andean mountain building did not preclude dispersal of
1066 lowland epiphytic orchids in the Neotropics. *Scientific Reports* 7: 4919.

1067 Pérez-Escobar, O. A., E. Lucas, C. Jaramillo, A. Monro, S. K. Morris, D. Bogarín, D. Greer, et
1068 al. 2019. The origin and diversification of the hyperdiverse flora in the Chocó biogeographic
1069 region. *Frontiers in Plant Science* 10: 1328.

1070 Pirie, M. D., P. J. M. Maas, R. A. Wilschut, H. Melchers-Sharrott, and L. W. Chatrou. 2018.
1071 Parallel diversifications of *Cremastosperma* and *Mosannona* (Annonaceae), tropical
1072 rainforest trees tracking Neogene upheaval of South America. *Royal Society open science*
1073 5: 171561.

1074 Pitman, N. C. A., P. N. Vargas, and J. W. Terborgh. 2017. Árboles comunes de los bosques
1075 inundados de Madre de Dios. *Biodiversidad Amazónica Vol 5* 1.

1076 Pyle, M. M., and R. P. Adams. 1989. In situ preservation of DNA in plant specimens. *Taxon* 38:
1077 576–581.

1078 Qiu, Y.-L., L. Li, T. A. Hendry, R. Li, D. W. Taylor, M. J. Issa, A. J. Ronen, et al. 2006.
1079 Reconstructing the basal angiosperm phylogeny: evaluating information content of
1080 mitochondrial genes. *Taxon* 55: 837–856.

1081 Queenborough, S. A., D. F. R. P. Burslem, N. C. Garwood, and R. Valencia. 2007a.
1082 Determinants of biased sex ratios and inter-sex costs of reproduction in dioecious tropical
1083 forest trees. *American Journal of Botany* 94: 67–78.

1084 Queenborough, S. A., D. F. R. P. Burslem, N. C. Garwood, and R. Valencia. 2007b. Habitat
1085 niche partitioning by 16 species of Myristicaceae in Amazonian Ecuador. *Plant Ecology*
1086 192: 193–207.

1087 Queenborough, S. A., D. F. R. P. Burslem, N. C. Garwood, and R. Valencia. 2007c.
1088 Neighborhood and community interactions determine the spatial pattern of tropical tree
1089 seedling survival. *Ecology* 88: 2248–2258.

1090 Rabeler, R. K., H. T. Svoboda, B. Thiers, L. A. Prather, J. A. Macklin, L. P. Lagomarsino, L. C.
1091 Majure, and C. J. Ferguson. 2019. Herbarium practices and ethics, III. *Systematic Botany*
1092 44: 7–13.

1093 Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other
1094 things). *Methods in Ecology and Evolution* 3: 217–223.

1095 Ripperger, S. P., M. Tschapka, E. K. V. Kalko, B. Rodriguez-Herrera, and F. Mayer. 2013. Life
1096 in a mosaic landscape: anthropogenic habitat fragmentation affects genetic population
1097 structure in a frugivorous bat species. *Conservation Genetics* 14: 925–934.

1098 Russo, S. E. 2003. Responses of dispersal agents to tree and fruit traits in *Virola calophylla*
1099 (Myristicaceae): implications for selection. *Oecologia* 136: 80–87.

1100 Rymer, P. D., C. W. Dick, G. G. Vendramin, A. Buonamici, and D. Boshier. 2013. Recent
1101 phylogeographic structure in a widespread 'weedy' Neotropical tree species, *Cordia*
1102 *alliodora* (Boraginaceae): Phylogeography of a widespread tropical tree species. *Journal of*
1103 *Biogeography* 40: 693–706.

1104 Sandoval-H, J., J. P. Gómez, and C. D. Cadena. 2017. Is the largest river valley west of the
1105 Andes a driver of diversification in Neotropical lowland birds? *The Auk: Ornithological*
1106 *Advances* 134: 168–180.

1107 Santamaría-Aguilar, D., R. Aguilar, and L. P. Lagomarsino. 2019. A taxonomic synopsis of
1108 *Virola* (Myristicaceae) in Mesoamerica, including six new species. *PhytoKeys* 134: 1–82.

1109 Santamaría-Aguilar, D., J. E. Jiménez, and R. Aguilar. 2019. *Otoba vespertilio* (Myristicaceae),
1110 una especie nueva de Mesoamérica. *Brittonia* 71: 369–380.

1111 Särkinen, T., R. T. Pennington, M. Lavin, M. F. Simon, and C. E. Hughes. 2012. Evolutionary
1112 islands in the Andes: persistence and isolation explain high endemism in Andean dry
1113 tropical forests. *Journal of Biogeography* 39: 884–900.

1114 Särkinen, T., M. Staats, J. E. Richardson, R. S. Cowan, and F. T. Bakker. 2012. How to open
1115 the treasure chest? Optimising DNA extraction from herbarium specimens. *PloS One* 7:
1116 e43808.

1117 Sauquet, H., J. A. Doyle, T. Scharaschkin, T. Borsch, K. W. Hilu, L. W. Chatrou, and A. Le
1118 Thomas. 2003. Phylogenetic analysis of Magnoliales and Myristicaceae based on multiple
1119 data sets: implications for character evolution. *Botanical Journal of the Linnean Society*
1120 142: 125–186.

1121 Sauquet, H., and A. Le Thomas. 2003. Pollen diversity and evolution in Myristicaceae
1122 (Magnoliales). *International Journal of Plant Sciences* 164: 613–628.

1123 Schley, R. J., M. de la Estrella, O. A. Pérez-Escobar, A. Bruneau, T. Barraclough, F. Forest, and
1124 B. Klitgård. 2018. Is Amazonia a 'museum' for Neotropical trees? The evolution of the
1125 *Brownea* clade (Detarioideae, Leguminosae). *Molecular Phylogenetics and Evolution* 126:
1126 279–292.

1127 Sharma, M. V., and J. E. Armstrong. 2013. Pollination of *Myristica* and other nutmegs in natural
1128 populations. *Tropical Conservation Science* 6: 595–607.

1129 Shaw, J., E. B. Lickey, J. T. Beck, S. B. Farmer, W. Liu, J. Miller, K. C. Siripun, et al. 2005. The
1130 tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for
1131 phylogenetic analysis. *American Journal of Botany* 92: 142–166.

1132 Shee, Z. Q., D. G. Frodin, R. Cámara-Leret, and L. Pokorny. 2020. Reconstructing the complex
1133 evolutionary history of the Papuasian *Schefflera* radiation through herbariomics. *Frontiers in*
1134 *Plant Science* 11: 258.

1135 Sinnott-Armstrong, M. A., A. E. Downie, S. Federman, A. Valido, P. Jordano, and M. J.
1136 Donoghue. 2018. Global geographic patterns in the colours and sizes of animal-dispersed
1137 fruits. *Global Ecology and Biogeography* 27: 1339–1351.

1138 Soltis, D. E., M. A. Gitzendanner, and P. S. Soltis. 2007. A 567-taxon data set for angiosperms:
1139 The challenges posed by Bayesian analyses of large data sets. *International Journal of*
1140 *Plant sciences* 168: 137–157.

1141 Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
1142 large phylogenies. *Bioinformatics* 30: 1312–1313.

1143 ter Steege, H., N. C. A. Pitman, O. L. Phillips, J. Chave, D. Sabatier, A. Duque, J.-F. Molino, et
1144 al. 2006. Continental-scale patterns of canopy tree composition and function across
1145 Amazonia. *Nature* 443: 444–447.

1146 Štorchová, H., R. Hrdličková, J. Chrtek Jr., M. Tetera, D. Fitze, and J. Fehrer. 2000. An
1147 improved method of DNA isolation from plants collected in the field and conserved in
1148 saturated NaCl/CTAB solution. *Taxon* 49: 79–84.

1149 Struth, L., J. Babault, and A. Teixell. 2015. Drainage reorganization during mountain building in
1150 the river system of the Eastern Cordillera of the Colombian Andes. *Geomorphology* 250:
1151 370–383.

1152 Swamy, V. 2017. Forest composition and spatial patterns across a Western Amazonian River
1153 basin: The influence of plant-animal interactions. *In* R. W. Myster [ed.], Forest structure,
1154 function and dynamics in Western Amazonia, 159–180. John Wiley & Sons, Ltd,
1155 Chichester, UK.

1156 Terborgh, J., P. Alvarez-Loayza, K. Dexter, F. Cornejo, and C. Carrasco. 2011. Decomposing
1157 dispersal limitation: limits on fecundity or seed distribution? *The Journal of Ecology* 99:
1158 935–944.

1159 Turchetto-Zolet, A. C., F. Cruz, G. G. Vendramin, M. F. Simon, F. Salgueiro, M. Margis-
1160 Pinheiro, and R. Margis. 2012. Large-scale phylogeography of the disjunct Neotropical tree
1161 species *Schizolobium parahyba* (Fabaceae-Caesalpinioideae). *Molecular Phylogenetics*
1162 *and Evolution* 65: 174–182.

1163 Valderrama, E., C. Sass, M. Pinilla-Vargas, D. Skinner, P. J. M. Maas, H. Maas-van de Kamer,
1164 J. B. Landis, et al. 2020. Unraveling the spiraling radiation: A phylogenomic analysis of
1165 Neotropical *Costus* L. *Frontiers in Plant Science* 11: 1195.

1166 Van Andel, T., M. A. Veltman, A. Bertin, H. Maat, T. Polime, D. Hille Ris Lambers, J. Tjoe Awie,
1167 et al. 2019. Hidden rice diversity in the Guianas. *Frontiers in Plant Science* 10: 1161.

1168 Wheelwright, N. T., and C. H. Janson. 1985. Colors of fruit displays of bird-dispersed plants in
1169 two tropical forests. *The American Naturalist* 126: 777–799.

1170 Wittmann, F., J. Schongart, J. C. Montero, T. Motzer, W. J. Junk, M. T. F. Piedade, H. L.
1171 Queiroz, and M. Worbes. 2006. Tree species composition and diversity gradients in white-
1172 water forests across the Amazon Basin. *Journal of Biogeography* 33: 1334–1347.

1173 Yu, Y., A. J. Harris, C. Blair, and X. He. 2015. RASP (Reconstruct Ancestral State in
1174 Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and Evolution* 87:
1175 46–49.

1176 Zhang, C., M. Rabiee, E. Sayyari, and S. Mirarab. 2018. ASTRAL-III: polynomial time species
1177 tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19: 153.

1178

1179

1180
1181

1182 **Table 1.** Summary of sequence data recovered for each sample for targeted and off-target
1183 (cpDNA) loci. Asterisks at sample names indicate samples that were not included in any
1184 phylogenetic analysis. Samples for which there was insufficient cpDNA data to include in
1185 phylogenetic analyses are marked with “n/a”.
1186

Sample	Avgerage sequence length	number of loci	ungapped bp cpDNA	Collection Year	Annual precipitation (mm)
***Otoba acuminata*_MM2720**	209.25	4	39,553	2001	3,021
***Otoba cyclobasis*_MT281**	183.9	73	3,399	1993	2,265
***Otoba glycycarpa*_LV25198**	227.88	96	32,401	2013	1,842
***Otoba gordoniifolia*_RZ196**	206.51	61	20,318	1996	1,933
Otoba_gracilipes_DC884*	192	2	n/a	1987	3,065
***Otoba latialata*_RC4751**	244.39	142	42,374	1987	2,610
***Otoba novogranatensis*_AG476**	102	1	3,533	1993	3,048
***Otoba novogranatensis*_CK681**	174	2	13,653	1988	3,499
Otoba novogranatensis*_EB500	208.91	11	n/a	1988	3,524
***Otoba novogranatensis*_GP2325**	186.75	4	41,378	1992	2,057
***Otoba novogranatensis*_LG20482**	299.92	157	111,863	1983	3,244
***Otoba novogranatensis*_WP16081**	90	1	47,902	1993	2,326
***Otoba novogranatensis*_WS36336**	457.56	217	n/a	2015	1,864
Otoba parvifolia*_DN9151	183	4	n/a	1989	3,870
***Otoba parvifolia*_MN37243**	173.47	17	28,315	1988	1,449
***Otoba parvifolia*_MS1182**	139.5	6	72,683	1995	2,109
***Otoba parvifolia*_RV19070**	215.62	86	n/a	1994	2,095

Otoba sp. nov._JP16902*	145.41	17	n/a	1992	2,438
Otoba sp. nov_RC5752*	126	5	n/a	1987	3,732
***Otoba vespertilio*_GM12543**	131	3	1,532	1988	3,561

1187
1188
1189

Table 2. Divergence times (Ma) for major clades of Otoba estimated from penalized likelihood analyses based on different calibrations from Magallon et al. (2015) and Massoni et al. (2015).

	Magallon et al. (2015)						Massoni et al. (2015)					
	UCLN median		UCLN min		UCLN max		UCLN mean		UCLN min		UCLN max	
	calibration	estimated age	calibration	estimated age	calibration	estimated age	calibration	estimated age	calibration	estimated age	calibration	estimated age
Magnioliales + Laurales	127.70	—	121.80	—	131.77	—	134.15	—	121.34	—	161.65	—
Laurales	114.90	—	109.07	—	120.64	—	122.16	—	112.05	—	145.59	—
Magnoliales	109.59	—	108.14	—	112.32	—	123.86	—	114.75	—	145.66	—
Myristicaceae	—	19.09	—	18.55	—	19.66	29.07	—	14.70	—	51.51	—
Otoba	—	7.28	—	7.04	—	7.51	—	8.69	—	6.57	—	11.36
So. Am. *O. novogranatensis* clade	—	5.82	—	5.62	—	6.00	—	6.60	—	5.47	—	8.23
Eastern Andean/Amazonian + Western Andean/Cen. Am. clade	—	6.54	—	6.32	—	6.75	—	7.67	—	5.98	—	9.83
Eastern Andean/Amazion clade	—	4.13	—	3.99	—	4.27	—	4.57	—	4.01	—	5.61
Western Andean/Central American clade	—	6.21	—	6.00	—	6.41	—	7.25	—	5.71	—	9.26

1195