A dated phylogeny of the papilionoid legume genus Canavalia reveals recent diversification by a pantropical liana lineage

Authors: Cristiane Snak, Mohammad Vatanparast, Christian Silva, Gwilym Peter Lewis, Matt Lavin, Tadashi Kajita, & Luciano Paganucci de Queiroz

Snak, Cristiane , Mohammad Vatanparast, Christian Silva, Gwilym Peter Lewis, Matt Lavin , Tadashi Kajita, and Luciano Paganucci de Queiroz. "A dated phylogeny of the papilionoid legume genus Canavalia reveals recent diversification by a pantropical liana lineage." Molecular Phylogenetics and Evolution 98 (May 2016): 133-146. DOI: 10.1016/j.ympev.2016.02.001.

A dated phylogeny of the papilionoid legume genus Canavalia reveals recent diversification by a pantropical liana lineage

Cristiane Snak, Mohammad Vatanparast, Christian Silva, Gwilym Peter Lewis, Matt Lavin, Tadashi Kajita, & Luciano Paganucci de Queiroz

[a] *Universidade Estadual de Feira de Santana, Departamento de Ciências Biológicas, Av. Transnordestina, s/n, Novo Horizonte, Feira de Santana 44036-900, Bahia, Brazil*
[b] *Department of Botany, National Museum of Natural History, MRC 166, Smithsonian Institution, Washington D.C. 20013-7012, USA*
[c] *Department of Biology, Graduate School of Science, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8-522, Japan*
[d] *Comparative Plant and Fungal Biology Department, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, United Kingdom*
[e] *Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman 59717, USA*
[f] *Iriomote Station, Tropical Biosphere Research Center, University of the Ryukyus, Uehara, Takketomi-cho, Okinawa 907-1541, Japan*

Abstract

Canavalia is a pantropical legume genus of lianas comprising approximately 60 species distributed in a wide range of habitats. In the last taxonomic revision, the genus was divided into four subgenera: *Canavalia* (Pantropical), *Catodonia* (Neotropical, excepting one species also found in the Old World), *Maunaloa* (Hawaiian), and *Wenderothia* (Neotropical). In this study, we reconstructed the phylogeny of *Canavalia* using a broad taxon sampling and analyses of nuclear (ETS and ITS) and plastid markers (*trnK*/*matK*). We evaluated the infrageneric classification of the genus and investigated its biogeographical history using molecular dating analyses and ancestral area reconstructions. The phylogenetic analyses resolved subgenus *Wenderothia* as monophyletic. Subgenus *Catodonia* needs to be recircumscribed and the relationships between subgenera *Canavalia* and *Maunaloa* remain unclear. *Canavalia* arose during the Miocene with a mean stem age estimate of 13.8 Ma and mean crown age estimate of 8.7 Ma, and most extant species evolved during the Pleistocene. Several climatic and geological events are chronologically coincident with the divergence of the major clades of *Canavalia* (glacial/interglacial periods, Andes uplift and the formation of Pebas and post-Pebas systems, closure of the Isthmus of Panama, and change in the direction of ocean currents). Ancestral area reconstructions for the early divergence of the genus are equivocal, although, some evidence suggests *Canavalia* originated in the wet forests of South America and achieved its current pantropical distribution through recent transoceanic dispersal. The evolution of *Canavalia* is better explained by a series of several processes than by discrete historical events.

1. Introduction

Diversification studies of plants using time-calibrated molecular phylogenies reveal a wide range of rate heterogeneity through time, among lineages, and among different geographical and ecological settings (Hughes et al., 2015). The patters of diversification in wet forests, seasonally dry forests, savannas and montane habitats differ in the timing of major radiations, net diversification rates, and species persistence over time (Pennington et al., 2004; Drummond, 2008; Simon et al., 2009; Hoorn et al., 2010; Särkinen et al., 2012; Simon and Pennington, 2012; Lohmann et al., 2013; Côrtes et al., 2015; Koenen et al., 2015). These highly variable explanations of the biodiversity among co-existing lineages suggest that unique historical events are not at play in shaping present day patterns of plant diversity.

The same might be true for explaining the diversity of individual lineages, such as that of the legume genus *Canavalia* Adans., the focus of our study. The pantropical genus *Canavalia* includes about 60 species of lianas in a wide range of habitats, including both wet and dry forests, savannas, and beach vegetation (Sauer, 1964; Aymard and Cuello, 1991). In addition to the lianescent habit, the genus is also characterized by trifoliolate leaves, a pseudoracemose inflorescence, and resupinate flowers with a bilabiate calyx where the lower (carinal) lip bears three small teeth and the upper

(vexillary) lip two larger lobes (Sauer, 1964). Some species have buoyant and impermeable seeds, which can drift for long periods over great distances (Sauer, 1964). Some species are used for food, cover crops, green manure, as ornamentals, and for medicine, e.g., jack bean [*Canavalia ensiformis* (L.) DC.] and sword bean [*C. gladiata* (Jacq.) DC.], both unknown in natural habitats (Sauer, 1964; NAS, 1979; Ekanayake et al., 2007).

The four subgenera of *Canavalia* (Sauer, 1964) show some ecological and biogeographical trends. Subgenera *Catodonia* J.D. Sauer and *Wenderothia* (Schltdl.) J.D. Sauer are restricted to the New World, except for *C. bonariensis* Lindl., which occurs in South America and Africa. Subgenus *Catodonia* is mainly from rainforest habitats, although some species pioneer beach vegetation. Subgenus *Wenderothia* occupies wet and seasonally dry forests and savannas. Subgenus *Canavalia* has a pantropical distribution and most of its species occur in wet forests, although some also pioneer beach vegetation (Sauer, 1964). Subgenus *Maunaloa* J.D. Sauer is endemic to the Hawaiian Islands and occupies wet and dry habitats (Sauer, 1964). *Canavalia* belongs to the tribe Diocleae (Queiroz et al., 2015), one of the groups included in the Canavanine Accumulating Clade (or the non-protein aminoacid accumulating clade), which represents the largest radiation of papilionoid legumes (Wojciechowski et al., 2004; Cardoso et al., 2012, 2013; LPWG, 2013). Phylogenetic studies of *Canavalia* based on morphological or molecular data have resolved the genus as monophyletic (Queiroz et al., 2003, 2015; Varela et al., 2004; Vatanparast et al., 2011). Morphology-based phylogenetic analyses of the Diocleae recovered *Canavalia* in a clade along with *Camptosema* W.J. Hook. & Arn., *Cleobulia* Mart. ex Benth., and *Cratylia* Mart. ex Benth (Queiroz et al., 2003). However, the studies using molecular markers indicated that the genus is sister to a clade that includes the remaining Diocleae (Doyle and Doyle, 1993; Kajita et al., 2001; Varela et al., 2004; Queiroz et al., 2015). Insufficient taxon sampling precluded the evaluation of phylogenetic relationships among the *Canavalia* subgenera with confidence.

Using Eocene and Miocene fossils supposedly belonging to different subgenera of *Canavalia*, Sauer (1964) hypothesized that the evolution of the genus was already advanced in the early Tertiary and thus the genus would have diverged from other Phaseoloid legumes during the Cretaceous. Sauer also suggested that *Wenderothia* could be the most primitive subgenus from which subgenera *Canavalia* and *Catodonia* would have been independently originated, and that subgenus *Maunaloa* would be an offshoot of subgenus *Canavalia*. Thus, Sauer (1964) inferred the monophyly of subgenera *Catodonia* and *Maunaloa* and the paraphyly of subgenera *Wenderothia* and *Canavalia*.

In this study, we conducted a time-calibrated molecular phylogenetic analysis of *Canavalia* using a broad taxonomic sampling that was designed to address the following questions: (1) Are the intuitive hypotheses of Sauer (1964) about the paraphyly of subgenera *Wenderothia* and *Canavalia*, and the monophyly of subgenera *Catodonia* and *Maunaloa* supported by a phylogenetic framework? (2) Where and when did *Canavalia* originate and diversify? (3) How many dispersal events occurred between the New and the Old World? (4) Are those dispersal events related to seed buoyancy?

2. Material and methods

2.1. Taxon sampling

The sampling includes 57 accessions of *Canavalia* corresponding to 47 of the 61 species currently recognized for the genus (17 of the 26 species of subg. *Canavalia*, 10/12 of subg. *Catodonia*, 3/6 of subg. *Maunaloa*, and 17/17 of subg. *Wenderothia*). Also sampled were five undescribed species (Snak, 2015; C. Snak, unpublished data). For *C. bonariensis* Lindl., *C. mattogrossensis* (Barb. Rodr.) Malme, and *C. villosa* Benth. we included two accessions that could represent different species based on distinct patterns of morphology and distribution. In this study, nine species of subg. *Canavalia* (*C. aurita* J.D. Sauer, *C. macrobotrys* Merr., *C. mollis* Wall. ex Wight & Arn., *C. papuana* Merr. & Perry, *C. raiateensis* J.W. Moore, *C. ramosii* J.D. Sauer, *C. regalis* Dunn, *C. sericea* A. Gray, *C. veillonii* I.C. Nielsen), two of subg. *Catodonia* (*C. mandibulata* J.D. Sauer, *C. sericophylla* Ducke), and three of subg. *Muanaloa* (*C. galeata* Gaudich., *C. kauaiensis* J.D. Sauer, *C. napaliensis* St. John) could not be sampled. *Cleobulia multiflora* Mart. ex Benth. and *Dioclea virgata* (Rich.) Amshoff were chosen as outgroups based on the phylogeny of Diocleae (Queiroz et al., 2015). A complete list of vouchers with associated GenBank accessions is presented in Table 1.

2.2. DNA extraction, amplification and sequencing

Total genomic DNA was extracted from silica gel-dried leaves using the 2× CTAB protocol of Doyle and Doyle (1987). For herbarium samples (ca. 60%), DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN GmbH, Hilden, Germany). Three DNA regions were selected for this study: plastid *trnK/matK* (the *matK* gene and partial flanking *trnK* introns), ribosomal nuclear ETS (partial 5′ end of the 18S ribosomal RNA gene and part of the External Transcribed Spacer) and ITS [3′ end of nuclear ribosomal 18S, 5.8S and flanking Internal Transcribed Spacers 1 and 2, and the 5′ end of 26S] (Table 2). We chose these regions because they have provided excellent resolution at different taxonomic levels across different clades of papilionoids (Queiroz et al., 2010; Queiroz and Lavin, 2011; Silva et al., 2012; Cardoso et al., 2012, 2013, 2015; Perez et al., 2013; Souza et al., 2013).

PCR reactions were performed using the TopTaq Master Mix Kit (QIAGEN GmbH, Hilden, Germany) according to the manufacturer's protocol, for a final volume of 10 μL. For herbarium samples, PCR reactions also included 2 μL of TBT-PAR [trealose, bovine serum albumin (BSA), polysorbate-20 (Tween-20)] (Samarakoon et al., 2013), and for ITS they also included 0.2 μL of DMSO 99.5% (dimethyl sulfoxide) in order to avoid secondary conformations (Table 2).

PCR products were cleaned using PEG 11% (Paithankar and Prasad, 1991), and then sequenced in both directions using the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Austin, Texas, USA) according to the following protocol: a hot start followed by 3 min of initial denaturation at 96 °C, 30 cycles of 96 °C denaturation for 20 s, 50 °C annealing for 15 s, and 60 °C extension for 4 min. Sequencing products were cleaned using isopropanol 80% and ethanol 70%, and analyzed on a 3130*xl* Genetic Analyzer (Applied Biosystems/HITACHI, Tokyo, Japan) at the Laboratório de Sistemática Molecular de Plantas of the Universidade Estadual de Feira de Santana (LAMOL/UEFS).

2.3. Alignment and phylogenetic analyses

The electropherograms were assembled and edited using the Geneious platform (Drummond et al., 2012a). Multiple sequence alignment was performed using MUSCLE (Edgar, 2004) with default settings. Alignments were inspected and adjusted manually using Geneious (Drummond et al., 2012a). The alignments for each data set are available in TreeBase (submission 18801). Unambiguous indels were coded for all data sets using the simple indel coding criteria (Simmons and Ochoterena, 2000), as implemented in the Seqstate software (Müller, 2005a). We performed maximum parsimony (MP) and Bayesian analyses (BA) for individual, nuclear (concatenated ETS and ITS), and combined (all regions) data sets. Conflicts among data sets were evaluated by the incongruence

Table 1
Voucher information and GenBank accession numbers for DNA sequences used in this study. All sequences of *Canavalia* were generated in this study. The sequences of outgroups are from the study of Queiroz et al. (2015). Classification according to Sauer (1964).

Taxon	Locality	Voucher	GenBank accession numbers		
			ETS	ITS	*trnK*/*matK*
Outgroups					
Muellera campestris (Mart. ex Benth.) M.J. Silva & A.M.G. Azevedo	Brazil, Bahia	Queiroz 13959 (HUEFS)	KC779808	–	KC779550
Galactia remansoana Harms	Brazil, Bahia	Queiroz 14659 (HUEFS)	KC779901	KC779782	KC779617
Cratylia mollis Mart. ex Benth.	Brazil, Bahia	Queiroz 8024 (HUEFS)	KC779879	KC779675	KC779568
Dioclea edulis Kuhlm.	Brazil, Bahia	Queiroz 15226 (HUEFS)	KC779835	KC779683	KC779573
Dioclea virgata (Rich.) Amshoff	Brazil, Rondônia	Cardoso 2917 (HUEFS)	KC779827	KC779723	KC779593
Cleobulia multiflora Mart. ex Benth.	Brazil, Bahia	Jesus 13 (HUEFS)	KC779819	KC779673	KC779564
Canavalia* subgenus *Canavalia					
C. africana Dunn	Togo, Tsévié	Ern2320 (B)	KT751371	KT751420	KT751466
C. boliviana Piper	Bolivia, La Paz	Siedel 7779 (HUEFS)	–	KT751425	KT751471
C. brasiliensis Mart. ex Benth.	Brazil, Maranhão	Snak 954 (HUEFS)	KT751377	KT751428	KT751474
C. campylocarpa Piper	Saint Lucia, Castries	Slane 341 (GH)	–	KT751429	KT751475
C. cathartica Thouars	Papua New Guinea, Mussau	Koie 1277 (B)	–	–	KT751477
C. dictyota Piper	Panama, Barro Colorado	Foster 1477 (F)	KT751380	KT751432	KT751479
C. ensiformis (L.) DC.	Cultivated	Snak 918 (HUEFS)	–	KT751435	KT751481
C. gladiata (Jacq.) DC.	Cultivated	Snak 1083 (HUEFS)	KT751386	KT751438	KT751484
C. gladiolata J.D. Sauer	Cultivated	Snak 1071 (HUEFS)	KT751387	KT751439	KT751485
C. lineata (Thunb.) DC.	Cultivated	Snak 1074 (HUEFS)	KT751391	KT751442	KT751490
C. madagascariensis J.D. Sauer	Madagascar, Antsiranana	Du Puy 239 (P)	–	KT751444	KT751492
C. megalantha Merr.	Mariana-Bonin Islands	Fosberg 35275 (P)	KT751397	KT751447	KT751495
C. piperi Killip & J.F. Macbr.	Brazil, Goiás	Coradin 7581 (HUEFS)	KT751404	KT751453	KT751502
C. plagiosperma Piper	Trinidad and Tobago	Johnson 1378 (BM)	KT751405	KT751454	KT751503
C. rosea (Sw.) DC.	Brazil, Rio de Janeiro	Snak 1015 (HUEFS)	KT751407	–	KT751505
C. vitiensis J.D. Sauer	Fiji, Viti Levu	Smith 4053 (P)	KT751417	KT751463	KT751514
Canavalia* subgenus *Catodonia					
C. beniensis Aymard & Cuello	Bolivia, La Paz	Beck 13328 (M)	KT751373	KT751423	KT751469
C. bonariensis Lindl.	Bolivia, Cochabamba	Wood 20957 (K)	KT751376	KT751427	KT751473
C. bonariensis Lindl.	Brazil, Paraná	Snak 518 (HUEFS)	KT751375	KT751426	KT751472
C. cassidea G.P. Lewis	Brazil, Bahia	Snak 1070 (HUEFS)	KT751378	KT751430	KT751476
C. dolichothyrsa G.P. Lewis	Brazil, Bahia	Snak 1054 (HUEFS)	KT751381	KT751433	KT751480
C. eurycarpa Piper	Bolivia, Santa Cruz	Wood 19656 (K)	KT751383	KT751436	KT751482
C. macropleura Piper	Venezuela, Aragua	Pittier 12557 (M)	KT751392	KT751443	KT751491
C. nitida (Cav.) Piper	Dominican Republic, Barahona	Howard 12238 (BM)	KT751398	KT751448	KT751496
C. parviflora Benth.	Brazil, Bahia	Snak 1029 (HUEFS)	KT751402	KT751451	KT751500
C. reflexiflora C. Snak et al.	Brazil, Minas Gerais	Snak 1112 (HUEFS)	KT751406	KT751455	KT751504
C. wurdackii Aymard & Cuello	Peru, Amazonas	Wurdack 1107 (NY)	KT751418	KT751464	KT751515
C. sp. 1	Brazil, Espírito Santo	Folli 4822 (K)	KT751410	KT751457	KT751507
Canavalia* subgenus *Maunaloa					
C. hawaiiensis O. Deg. et al.	Hawaii, Kona	Picco 31530 (M)	–	–	KT751487
C. molokaiensis O. Deg. et al.	Hawaii	Hughes 43 (BISH)	–	HQ707545	HQ707545
C. pubescens Hook. & Arn.	Hawaii	Hobdy 1899 (BISH)	–	HQ707576	HQ707576
Canavalia* subgenus *Wenderothia					
C. acuminata Rose	Mexico, Oaxaca	Javier 860 (MBM)	KT751370	KT751419	KT751465
C. albiflora Ducke	Brazil, Maranhão	Snak 995 (HUEFS)	KT751372	KT751421	KT751467
C. altipendula (Piper) Standl.	Jamaica, Manchester	Howard 14707 (BM)	–	KT751422	KT751468
C. bicarinata Standl.	Colombia, Magdalena	Shultze-Kraft 8772 (HUEFS)	KT751374	KT751424	KT751470
C. concinna J.D. Sauer	Brazil, Mato Grosso	Borges 724 (HUEFS)	KT751379	KT751431	KT751478
C. dura J.D. Sauer	Mexico, Chiapas	Martinez 19412 (BM)	KT751382	KT751434	–
C. glabra (M. Martens & Galeotti) J.D. Sauer	Guatemala, San Marcos	Harmon 4726 (MO)	KT751385	–	–
C. grandiflora Benth.	Brazil, Mato Grosso	Borges 830 (HUEFS)	KT751388	KT751440	KT751486
C. hirsutissima J.D. Sauer	Mexico, Oaxaca	Sousa 10539 (BM)	KT751389	KT751441	KT751488
C. lenta Benth.	Brazil	Pohl 2948 (F)	KT751390	–	KT751489
C. mattogrossensis (Barb. Rodr.) Malme	Brazil, Mato Grosso do Sul	Snak 842 (HUEFS)	KT751394	KT751445	KT751493
C. mattogrossensis (Barb. Rodr.) Malme	Bolivia, Santa Cruz	Wood 24094 (K)	KT751395	KT751446	KT751494
C. matudae J.D. Sauer	Guatemala, San Marcos	Steyermark 36600 (F)	KT751396	–	–
C. obidensis Ducke	Brazil, Pará	Ducke 11810 (US)	KT751399	–	KT751497
C. oxyphylla Standl. & L.O. Williams	Costa Rica, Cartago	Dobbeler 1644 (BM)	KT751400	KT751449	KT751498
C. palmeri (Piper) Standl.	Mexico, Chiapas	Reyes-García 5804 (HUEFS)	KT751401	KT751450	KT751499
C. picta Killip & J.F. Macbr.	Brazil, Paraná	Snak 916 (HUEFS)	KT751403	KT751452	KT751501
C. rupicola Standl. & L.O. Williams	Mexico, Campeche	Martinez 31434 (BM)	KT751408	KT751456	KT751506
C. septentrionalis J.D. Sauer	Mexico, Tamaulipas	Stanford 1017 (MO)	KT751409	–	–
C. villosa Benth.	Costa Rica, Puntarenas	Zamora 2329 (K)	KT751416	KT751462	KT751513
C. villosa Benth.	Belize, Cayo	Peña 1058 (BM)	KT751415	KT751461	KT751512
C. sp. 2	Venezuela, Apure	Keller-Grein 18269 (HUEFS)	KT751411	KT751458	KT751508
C. sp. 3	Colombia, Comisaría del Putumayo	Cuatrecasas 11110 (US)	KT751412	–	KT751509
C. sp. 4	Ecuador, Ollerana	Whitefoord 801 (BM)	KT751413	KT751459	KT751510
C. sp. 5	Brazil, Goiás	Dawson 15030 (NY)	KT751414	KT751460	KT751511
Canavalia* subgenus *Canavalia*/*Maunaloa					
C. favieri I.C. Nielsen	New Caledonia, South Province	Veillon 7465 (P)	KT751384	KT751437	KT751483

Table 2
Sequence of the primers used for PCR amplification and sequencing, and PCR conditions.

DNA region	Primer name	Primer Sequence 5′–3′	References	PCR Conditions					
				Pre-melting	Denaturation (I)	Primer Annealing (II)	Primer Extension (III)	Cycles (I + II + III)	Final Extension
ETS	18S-IGS	GACACAAGCATATGACTACTGGCAGGATCAACCAG	Baldwin and Markos (1998)	94 °C (3 min)	94 °C (1 min)	55 °C (1 min)	72 °C (1.5 min)	30	72 °C (7 min)
	ETS-Dio	GCTTGTGCATCGAACGGTTGG	Queiroz et al. (2015)						
ITS	17SE (F)	ACG AAT TCA TGG TCC GGT GAA GTG TTC G	Sun et al. (1994)	94 °C (3 min)	94 °C (1 min)	52 °C (40 s)	72 °C (2.5 min)	28	72 °C (7 min)
	26SE (R)	TAG AAT TCC CCG GTT CGC TCG CCG TTA C	Sun et al. (1994)						
	SSF	GTCGTAACAAGGTTTCCGTAG	Kollipara et al. (1997)	Following manufacturer's protocol for sequencing					
	LSR	GTTAGTTTCTTTTCCTCC	Kollipara et al. (1997)						
trnK/matK	matK685F	GTATCGCACTATGTATTATTTGA	Wojciechowski et al. (2004)	94 °C (3 min)	94 °C (40 s)	55 °C (45 s)	72 °C (1 min)	36	72 °C (7 min)
	matK4La	CCTTCGATACTGGGTGAAAGAT	Wojciechowski et al. (2004)						
	matK1100L	TTCAGTCGTACCGAGTCAAATG	Wojciechowski et al. (2004)						
	matK4R	CATCTTTCACCCAGTAGCGAAG	Hu et al. (2000)						
	matK1932R	CAGACCGCTTACTAATGGG	Hu et al. (2000)						
	trnK2R	CCCGGAACTAGTCGGATG	Wojciechowski et al. (2004)						

length difference test (ILD; Farris et al., 1995), performed in PAUP* v.4.0b10 (Swofford, 2002) between nuclear regions and between the nuclear and plastid regions, using a heuristic search with 1000 replicates, random taxa-addition, and tree bisection and reconnection (TBR) branch-swapping, saving 15 trees per replicate.

Parsimony analyses were performed using PAUP* v.4.0b10 (Swofford, 2002) with Fitch parsimony as the optimality criterion (Fitch, 1971). Heuristic searches were performed with 1000 random taxa-addition replicates, and TBR branch-swapping, and retaining up to 15 trees per replicate. The resulting trees were used as starting trees for a subsequent round of TBR swapping. Clade support was evaluated using non-parametric bootstrapping (Felsenstein, 1985) with 2000 replicates (Hedges, 1992; Müller, 2005b), simple taxon-addition, and TBR algorithm, saving 15 trees per replicate. Only bootstrap percentages (BP) > 85 were considered as strong support (Kress et al., 2002).

Bayesian analyses were performed using MrBayes v.3.2.5 (Ronquist and Huelsenbeck, 2003) in the CIPRES Science Gateway v.3.3 (Miller et al., 2010). Nucleotide substitution models were selected using the Akaike information criterion (AIC) in MrModeltest v.2.3 (Nylander, 2004) according to each DNA region (Table 3). The coded indels were treated as a separate binary data set. Two runs using the Metropolis-coupled MCMC (Markov Chain Monte Carlo) algorithm, each with four random-initiated chains (one 'cold' and three 'heated') (Huelsenbeck et al., 2001), involved 10 million generations and these were sampled every 1000 generations. The convergence of the runs was assessed by checking if the standard deviation of split frequencies reached a value below 0.01. Trees generated before convergence were excluded as burnin (trnK/matK 1400 trees; ETS 1200 trees; ITS 840 trees; nuclear 1500 trees; combined 3600 trees) and the effective sample size (ESS) of all parameters was checked to verify if the values were >200. The remaining trees were summarized into a majority-rule consensus tree including the posterior probabilities (PP) as branch support estimates. Only PP values $\geqslant 95$ were considered as a strong support (Erixon et al., 2003). *Cleobulia multiflora* was chosen as the outgroup in the Bayesian analyses. The trees were edited using FigTree v.1.3.1 (Rambaut, 2009).

Due to the different relationships recovered among the major clades of *Canavalia* in the combined analyses, ITS analyses and MP analyses of the *trnK*/*matK* data set, and in the ETS analyses and MP analysis of the nuclear data set, we performed a test of alternative hypotheses using the Shimodaira–Hasegawa test (SH test; Shimodaira and Hasegawa, 1999). For that, we first generated unconstrained maximized likelihood estimate (MLE) trees for the combined data set under the model GTR + I + Γ using RAxML v.8.1.20 (Stamatakis, 2006). Thereafter, constrained MLE trees for each alternative hypothesis [constraint 1: Clade A sister to Clade (B + C); constraint 2: Clade C sister to Clade (A + B)] were also generated. Tests between unconstrained and constrained MLE topologies were performed in PAUP* v.4.0b10 (Swofford, 2002) using RELL and 1000 bootstrap replicates. The differences between the topologies were considered significant when $p < 0.05$.

2.4. Divergence time estimation

We carefully evaluated the illustrations and photographs of the fossils of *Canavalia* cited in the literature: *Canavalia eocenica* Berry, *C. acuminata* Berry (Berry, 1916, 1930), *Leguminosites andiraformis* Berry (Berry, 1930) from multiple Eocene formations, and *Leguminosites canavaliformis* Berry and *Canavalia miocenica* Berry from Miocene formations (Berry, 1925). These fossils represent only impressions of detached leaflets and none could be confidently ascribed to *Canavalia*. Thus, to date the phylogeny, we first estimated a time-calibrated phylogeny of the Canavanine Accumulating Clade of the Papilionoid legumes to estimate the ages of the

clades of our interest, and then we performed a secondary time calibration in the phylogeny of *Canavalia*.

2.4.1. Divergence time estimation in the Canavanine Accumulating Clade

Divergence time estimates for the Canavanine Accumulating Clade were obtained using a *matK* data set from the study of Queiroz et al. (2015). Species of the clades Andira, Dalbergioids s. l., and Baphioids were used as outgroups (Appendix A). We used the BEAST package v.1.8.2 (Drummond et al., 2012b) with a GTR substitution model and gamma distribution (GTR + Γ), an uncorrelated lognormal relaxed clock model, a tree prior with a Yule speciation model, and a random starting tree. We used six calibration points. Five of these were based on the mean ages derived from a previous analyses of the Leguminosae (Lavin et al., 2005): the Most Recent Common Ancestor (MRCA) of the Dalbergioids s.l. (node 1; mean = 55.3 Mya, stdev = 0.5), the Indigofereae (node 2; mean = 30 Mya, stdev = 1), the Millettioids and Phaseoloids (node 3; mean = 45.2 Mya, stdev = 0.3), the Millettioids s.l. (node 4; mean = 36.2 Mya, stdev = 0.5), and the Hologalegina Clade (node 5; mean = 50.6 Mya, stdev = 0.9), using a normal prior for each node. The sixth calibration point, derived from a set of fossil leaflets from the Puryear, Bolden, and Bovay Clay Pits from the Middle Eocene very similar to those of *Machaerium* Pers. (Herendeen, 1992), used to set the minimum age for the *Machaerium*, *Aeschynomene* L., and *Steinbachiella* Harms stem group using a lognormal prior distribution (node 6; mean = 3.69 Mya, log stdev = 0.005, offset = 0).

A BEAST file was generated in BEAUti v.1.8.2 (Drummond et al., 2012b). The major clades were preset based on Cardoso et al. (2013) and Queiroz et al. (2015), without forcing them to be monophyletic. Fifteen independent runs of MCMC chains were run for 10 million generations, sampling trees every 10,000 generations. These runs were performed using Beast in the CIPRES Science Gateway v.3.3 (Miller et al., 2010). The output files were visualized in Tracer v.1.6 (Rambaut and Drummond, 2013) to assess convergence of the runs and ESS values (only ESS > 200 were accepted). We discarded the first 350 trees from each run as burnin and combined the time tree files in LogCombiner v.1.8.2 (Drummond et al., 2012b). The results were summarized in a maximum clade credibility tree using TreeAnnotator v.1.8.2 (Drummond et al., 2012b). The chronogram was visualized and edited using FigTree v.1.3.1 (Rambaut, 2009).

2.4.2. Divergence time estimation in Canavalia

Divergence time estimates were assessed using a combined data set composed of *trnK*/*matK*, ETS and ITS sequences for 50 taxa of *Canavalia*, along with five species of Diocleae from the Galactia and Dioclea clades (sensu Queiroz et al., 2015). *Muellera campestris* (Mart. ex Benth.) M.J. Silva & A.M.G. Azevedo was selected as the outgroup for this analysis based on Queiroz et al. (2015). *Canavalia glabra*, *C. lenta*, *C. matudae*, *C. septentrionalis*, and *C.* sp3 were excluded from the analyses due to the missing data present in their sequences. To estimate divergence times in *Canavalia* we used the BEAST package v.1.8.2 (Drummond et al., 2012b) with a GTR substitution model and gamma distribution (GTR + Γ), an uncorrelated lognormal relaxed clock model, a tree prior with a Yule speciation model, and a random starting tree. The calibration was based on the mean ages derived from the previous analyses of the Canavanine Accumulating Clade for the MRCA of the Diocleae (mean = 13.6 Mya, stdev = 2) and *Canavalia* (mean = 8.4 Mya, stdev = 2), using a normal prior for each node. A BEAST file was generated in BEAUti v.1.8.2 (Drummond et al., 2012b) and the major clades of *Canavalia* were preset based on the topology obtained from the combined Bayesian analysis, without forcing them to be monophyletic. The analyses were performed according to the same procedures employed in the divergence time estimation for the Canavanine Accumulating Clade. However, only five independent runs of MCMC chains were run, and the first 100 trees were removed from each run as burnin. The nomenclature and ages for the geological time scale follow Cohen et al. (2013).

2.5. Biogeographical reconstructions

Species distributions and habitat data were compiled from herbarium specimens and literature (Sauer, 1964; Aymard and Cuello, 1991). From these data we established six biogeographic areas relevant to *Canavalia* (Fig. 2):

A. Northern and Northwestern South America: this corresponds to a huge area of wet forests of French Guiana, Suriname, Guyana, Venezuela, Colombia, Ecuador, and Brazilian, Peruvian and Bolivian Amazon.
B. South American dry diagonal: areas under markedly seasonal climate in Northeastern and Central Brazil, Northern and Central Argentina, Southern Bolivia, Western and Central Paraguay, and Western Peru.
C. Eastern South America: another area of wet forests in Eastern Brazil and Uruguay to Northeastern Argentina and Eastern Paraguay.
D. Mesoamerica: from Southeastern United States, Mexico, Belize, Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica, and Panama.
E. Old World: Africa, Asia, Oceania, and Hawaiian Islands.
F. Caribbean: Bahamas, and Greater and Lesser Antilles.

Each taxon was coded for presence/absence in the aforementioned six geographic areas (Appendix B). Ancestral areas of *Canavalia* were estimated using a maximum likelihood analysis of geographical range evolution using the package BioGeoBEARS 0.2.1 (Matzke, 2013) as implemented in R 3.2.2 (R Core Team, 2015). We tested a total of six models (Matzke, 2013): Dispersal–Extinction–Cladogenesis (DEC; Ree and Smith, 2008), Dispersal–Vicariance Analysis (DIVA; Ronquist, 1997), BayArea model (Landis et al., 2013), and an additional parameter (+J) for each model. This additional parameter simulates the process of founder-event speciation, which can be associated with long distance dispersal and subsequent divergence of colonizing lineages (Matzke, 2013). We used likelihood-ratio tests and AIC values to compare the fit of these models to the data. For this analysis, we used the BEAST maximum clade credibility tree without the outgroups and the cultivated species (*Canavalia ensiformis*, *C. gladiata* and *C. plagiosperma* Piper). The maximum number of ancestral areas to be reconstructed at each node was set to six.

2.6. Evolution of seed buoyancy

The evolution of the seed buoyancy within *Canavalia* was evaluated by reconstructing ancestral character states. The information about seed buoyancy was based on Sauer (1964) and personal observations. The reconstruction was performed in Mesquite v.3.02 (Maddison and Maddison, 2015) using parsimony and unordered character states. The seed buoyancy was treated as a discrete character with two states [absent (0), present (1)] (Appendix C). The analysis was based on a set of 1000 trees which correspond to the last 200 trees from each of the five MCMC runs of the divergence time estimation analysis, without the outgroups. The results were plotted in the BEAST maximum clade credibility tree obtained from the divergence time estimation analysis.

Table 3
Features of the DNA data sets used in this study based on one of the most parsimonious trees from the combined parsimony analysis, and nucleotide substitution models selected for the Bayesian analyses. (bp = base pairs; CI = consistency index; RI = retention index; ts:tv = transition/transversion ratio).

DNA region	Aligned length (bp)	Number variable sites		Number potentially parsimony informative sites		Number of changes/variable sites	Fitch tree length		CI	RI	ts:tv	Bayesian model
ETS region	415	193	(46.51%)	128	(30.84%)	1.78	345		0.73	0.90	1.97	GTR+I+G
ITS region	670	237	(35.52%)	166	(24.78%)	1.84	439		0.71	0.89	1.87	Mixed
ITS1	265	108	(40.75%)	78	(29.43%)	1.94	210		0.70	0.89	1.80	HKY+G
5.8S	164	8	(4.88%)	4	(2.44%)	1.50	12		0.67	0.56	All ts	K80+I
ITS2	241	122	(50.62%)	84	(34.85%)	1.78	217		0.71	0.89	1.78	HKY+G
18S/26S	185	8	(4.32%)	5	(2.70%)	1.88	15		0.53	0.77	6.50	K80+I
trnK introns	390	33	(8.46%)	23	(5.90%)	1.27	42		0.79	0.86	0.40	F81+I
matK gene	1533	104	(6.78%)	52	(3.39%)	1.26	131		0.82	0.89	0.82	Mixed
matK (1st positions)	511	32	(6.26%)	13	(2.54%)	1.28	41	(30.30%)	0.80	0.89	0.71	GTR+I
matK (2st positions)	511	27	(5.28%)	11	(2.15%)	1.15	31	(23.66%)	0.90	0.87	0.82	HKY+G
matK (3st positions)	511	45	(8.81%)	28	(5.48%)	1.31	59	(45.04%)	0.78	0.90	0.90	GTR+I
Gaps	102	102	(100.00%)	52	(50.98%)	1.29	132		0.77	0.90		Variable
Nuclear (ETS, ITS, 18S/26S)	1270	439	(34.57%)	299	(23.54%)	1.82	799		0.72	0.90		Mixed
Combined (all data)	3295	678	(20.58%)	426	(12.93%)	1.73	1104		0.75	0.89		Mixed

3. Results

3.1. Molecular data sets

We produced 146 new sequences for *Canavalia* (49 of ETS, 46 of ITS, and 51 of *trnK*/*matK*). Complete sequences were generated for all taxa, except for *C. altipendula*, *C. boliviana*, *C. campylocarpa*, *C. cathartica*, *C. ensiformis*, *C. hawaiiensis* and *C. madagascariensis*, for which we failed to amplify the ETS; *C. cathartica*, *C. glabra*, *C. hawaiiensis*, *C. lenta*, *C. matudae*, *C. obidensis*, *C. rosea*, *C. septentrionalis* and *C.* sp. 3, for which we failed to amplify the ITS; *C. dura*, *C. glabra*, *C. matudae* and *C. septentrionalis*, for which we failed to amplify the *trnK*/*matK*.

The most variable data set was ITS2, followed by ETS and ITS1, and the least variable data sets were the 18S/26S and the 5.8S. The ETS, ITS1, ITS2 and *matK* regions performed similarly and better than 5.8S and 18S/26S and *trnK* data sets [informativeness measured by the retention index (RI) of each data set]. The matrix containing the coded indels was composed of 28 gaps from ETS, 59 from ITS, and 15 from *trnK*/*matK* (Table 3).

3.2. Phylogenetic analyses

The ILD test indicated no incongruence between nuclear data sets ($p = 0.98$) or between nuclear and plastid data sets ($p = 0.28$). We did not detect strongly supported incongruent nodes. The individual and combined MP and BA analyses recovered the same major clades (Fig. 1, Appendix D). Three main clades were recovered with strong support within *Canavalia* (Fig. 1): Clade A was recovered as sister to the clades B and C, but this relationship was weakly supported. Alternative relationships among these three main clades were not rejected by the SH test (constraint 1: $p = 0.487$; constraint 2: $p = 0.275$).

Clade A comprises the species of *Canavalia* subg. *Catodonia* (except *C. macropleura*) together with the New Caledonian *C. favieri*. Within clade A, two clades were strongly supported in the BA and MP analyses, Clade A1, including *C. bonariensis*, *C. nitida*, and *C. favieri*, and Clade A2, including the remaining species of subgenus *Catodonia* which are confined to the wet forests of South America.

Clade B comprises the species *Canavalia macropleura*, currently ascribed to subgenus *Catodonia*, sister to Clade B1, which comprises the species of subgenera *Canavalia* and *Maunaloa*. Clade B1 is poorly resolved in the Bayesian 50% consensus tree and in the MP strict consensus tree, and most of the clades recovered were weakly supported.

Clade C comprises the species of subgenus *Wenderothia*. Three main clades were recovered within Clade C in the BA and MP combined analyses, however, the relationships among them are weakly supported. The clade C1 was resolved as sister of the remaining taxa of the clade in the BA combined analysis with weak support. However, its sister relationship was better supported in the BA analysis of the ITS data set (PP 0.95). The remaining taxa were recovered in a polytomy, which includes *C. altipendula* and two large clades. Although not well supported, these three lineages are geographically consistent: *C. altipendula* is the only species of subg. *Wenderothia* in the Caribbean Islands, Clade C2 includes Mesoamerican and Mexican species, and Clade C3 contains only South American species.

3.3. Divergence time estimation in the Canavanine Accumulating Clade

Divergence time estimates indicated an origin of the tribe Diocleae between the Upper Eocene and the Miocene, with a mean stem age of 33 Mya [95% highest posterior density (HPD), 35.6–30 Mya] and mean crown age of 13.6 Mya (95% HPD, 18.1–9.7 Mya) (nodes 7 and 8 respectively, Fig. 2). For *Canavalia* our analysis indicated a Miocene origin for the genus with a mean crown age of 8.4 Mya (95% HPD, 12.8–4.2 Mya) (node 9, Fig. 2). The full tree is available upon request to the first author.

3.4. Biogeography and divergence time estimation of Canavalia

Divergence time estimates indicated a Miocene origin of the genus *Canavalia* with a mean crown age of 8.7 Mya (95% HPD, 11.1–6.7 Mya) (Fig. 3). The major clades diverged between the Upper Miocene and the Pliocene (6.2–3.6 Mya), and the diversification of most extant species was estimated to be in the Pleistocene (in the last 2.5 Mya; Fig. 3; Table 4; Appendix E).

Ancestral range estimation recovered the DEC+*j* model as the best-fit model for the *Canavalia* data set ($\ln L = -101.72$, AICwt = 0.99) (Appendix F). This result suggests that long-distance dispersal may have had an important role in the history of *Canavalia*, as shown by the values obtained for the free parameters of the analysis ($d = 0.0258$, $e = 0$, and $j = 0.0342$).

The ancestral area reconstruction did not unambiguously identify where the MRCA of *Canavalia* first evolved (Table 4; node 1,

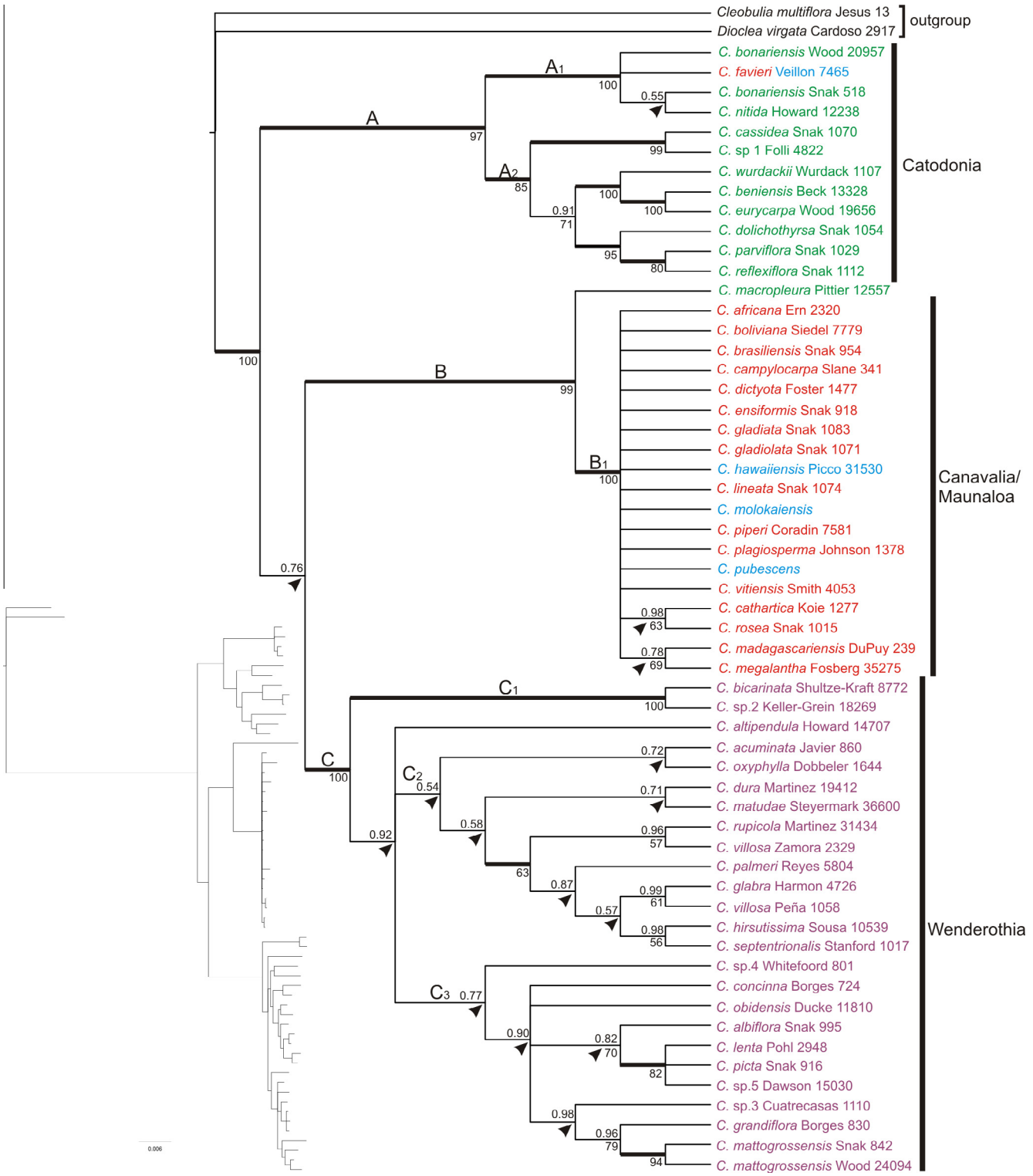

Fig. 1. Bayesian 50% consensus tree resulting from the combined nuclear (ETS, ITS) and plastid (*trnK*/*matK*). Bayesian posterior probabilities (only values >50%) and parsimony bootstrap support values are reported above and below branches, respectively. Branches supported by posterior probability 1 are in bold. Arrowheads indicate nodes collapsed in the Maximum Parsimony strict consensus. The colors on the tips indicate the subgenera according to Sauer's (1964) classification – blue: subg. *Maunaloa*; green: subg. *Catodonia*; purple: subg. *Wenderothia*; red: subg. *Canavalia*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 3; Appendix G). The divergence within Clade A was estimated to be in the Upper Miocene (6.2 Mya; node 2, Fig. 3), and its ancestor could have occurred in Eastern South America (40%). The expansion of the distribution of this ancestor to the Caribbean and Old World may represent the first long-distance dispersal event within the genus. The divergence within Clade B and Clade C was estimated to be in the Upper Miocene (7.6 Mya), and the MRCA was reconstructed from Northern and Northwestern South America (16%). This lineage experienced an expansion of its distribution to the South American dry diagonal, the Old World and the

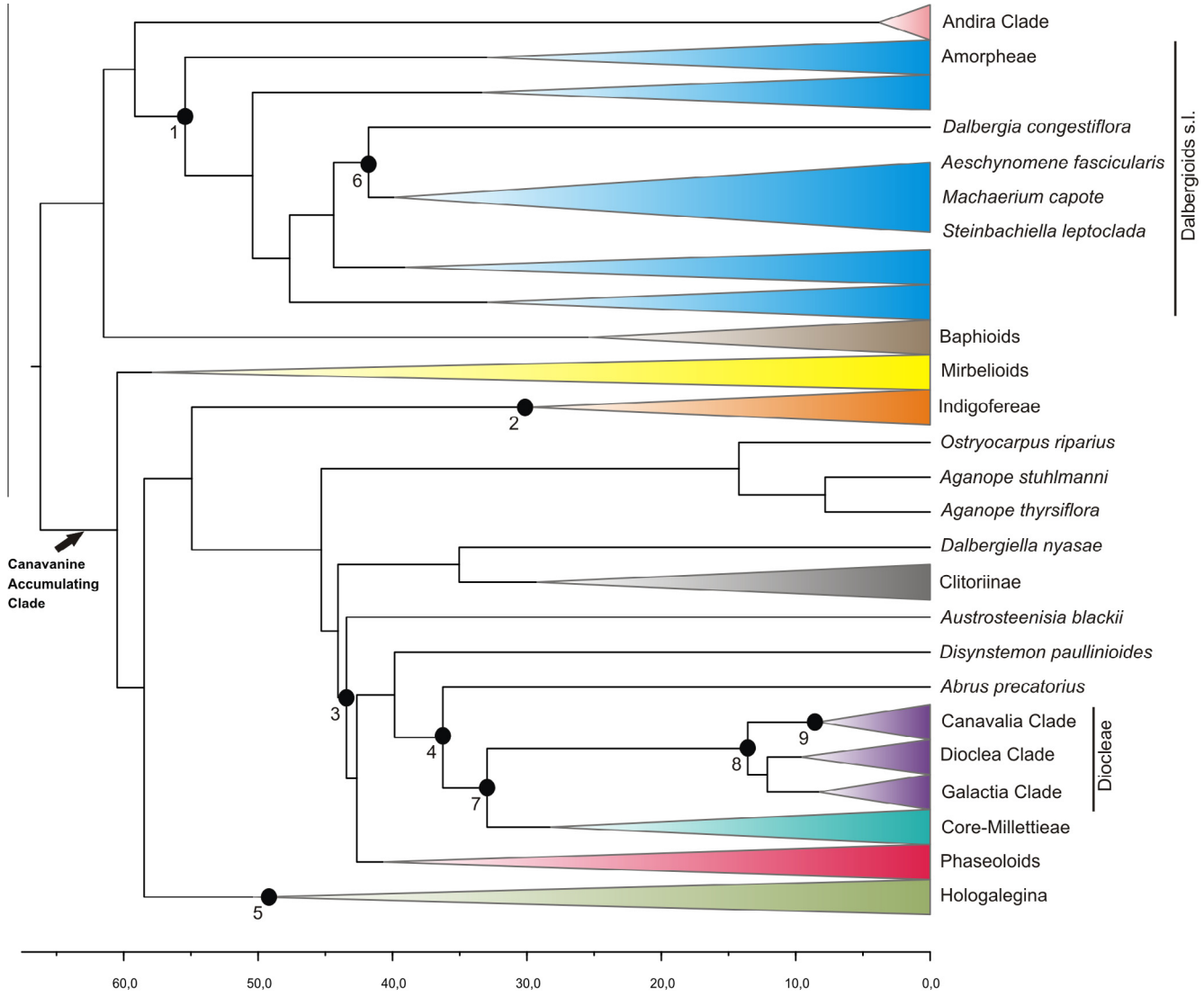

Fig. 2. Maximum clade credibility chronogram of the Canavanine Accumulating Clade obtained from the analysis of the *matK* data set of Queiroz et al. (2015). Black circles numbered from one to six refer to the calibration points. The time-scale is in millions of years ago (Mya).

Caribbean in the Pleistocene (1.3 Mya), in clade B1; and to the Mesoamerican, Caribbean, South American dry diagonal (more specifically to the Cerrado), and Eastern South America in the last 3.6 Mya in clade C (Table 4; Fig. 3).

At least two events of transoceanic dispersal from the New World to the Old World could have occurred, one in Clade A with respect to the dispersal of *Canavalia bonariensis* and the ancestor of *C. favieri*, and at least one in clade B1. However, the low resolution within clade B1 did not allow us to identify how many dispersal events could have occurred in this clade.

3.5. Evolution of seed buoyancy

The ancestral character state reconstruction for seed buoyancy is shown in Appendix H. The reconstructions indicated that seed buoyancy arose several times independently within *Canavalia*. In Clade A, three independent origins were estimated. Due to the low internal resolution, we could not confidently specify how many times buoyant seeds arose in Clade B. Unsurprisingly, no origins for seed buoyancy were estimated for clade C because none of its extant species possess buoyant seeds.

4. Discussion

4.1. Phylogenetic relationships within Canavalia

Previous phylogenetic studies which included *Canavalia* recovered distinct relationships among its subgenera. Studies based on the ITS region recovered subg. *Wenderothia* (corresponding to Clade C in Fig. 1) as sister to the remaining subgenera (Varela et al., 2004; Vatanparast et al., 2011) while analyses based on plastid data (*atpB–rbcL*, *ndhD–ndhE*, *trnH–psbA*, *rps16* intron, *trnD–trnT*, and *trnK/matK*) recovered a clade with subg. *Maunaloa* nested within subg. *Canavalia* (i.e., Clade B1 in Fig. 1) as sister to a clade including the remaining subgenera (Vatanparast et al., 2011). More recently, combined analyses of plastid and nuclear data supported subg. *Catodonia* (roughly corresponding to Clade A in Fig. 1) as sister to a clade including the remaining subgenera (Queiroz et al., 2015). The lack of support in the backbone of the tree from the multilocus analyses reported here and the results of the test of alternative hypotheses (Shimodaira–Hasegawa Test) did not assist in resolving potential relationships among the major clades of *Canavalia*.

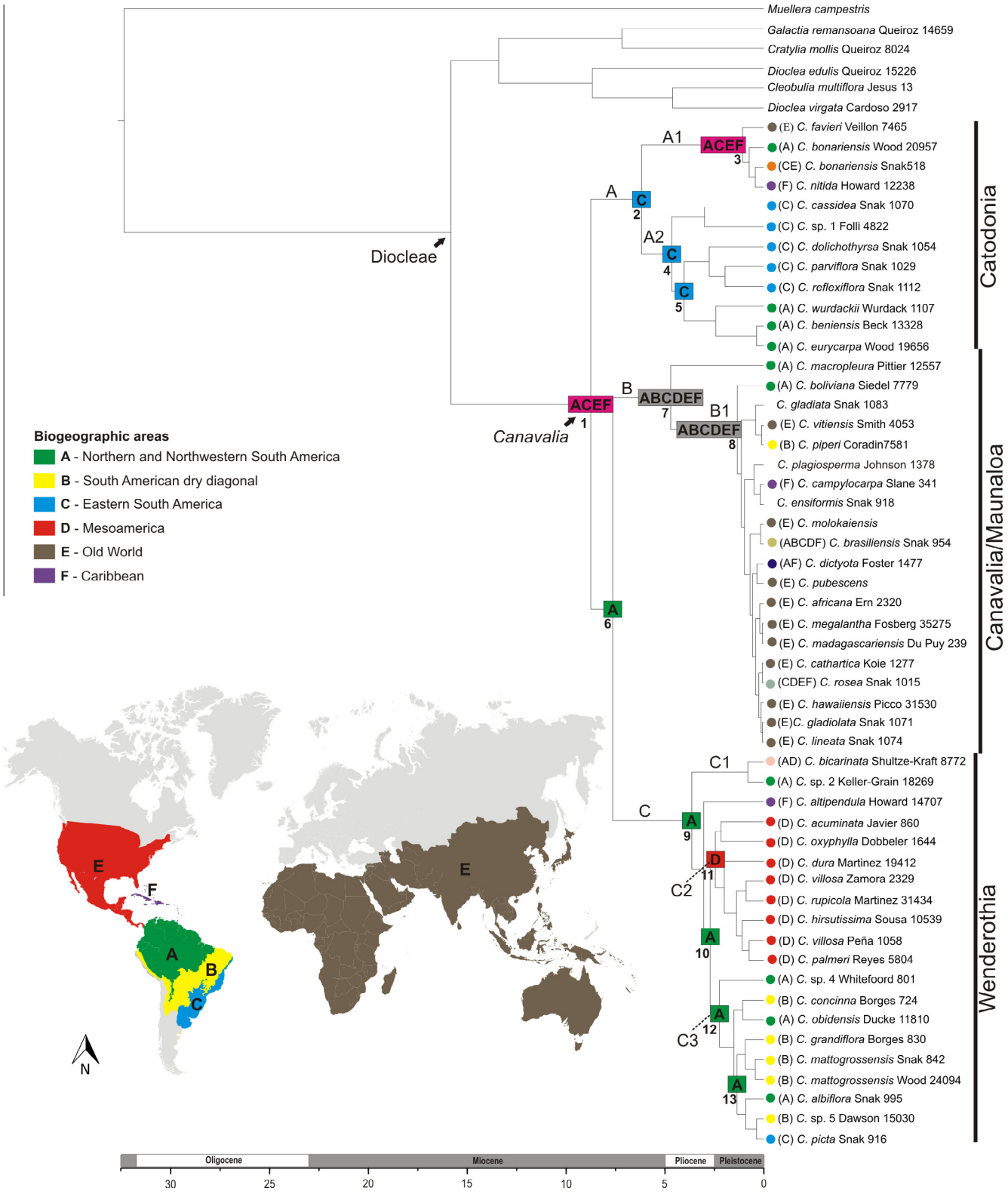

Fig. 3. Chronogram of *Canavalia* with ancestral reconstructions and tip states indicated. Internal colored squares indicate most likely ancestral area recovered by BioGeoBEARS under the DEC+J model. Colored circles behind tips indicate the character states of the sampled extant species. Black arrows indicate calibration points employed in the dating analyses. Geological eras are indicated in the time-scale, in millions of years ago (Mya). Numbers on the nodes refer to the clades listed in Table 4. The tree displayed here is the maximum clade credibility (MCC) tree from BEAST analyses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Among the four subgenera of *Canavalia* proposed by Sauer (1964), only subg. *Wenderothia* is unequivocally supported as monophyletic, thus rejecting Sauer's (1964) hypothesis that it represents a basal paraphyletic grade in genus *Canavalia*. Species of this subgenus have been grouped since Bentham (1837), who classified them based on the presence of rostrate keel petals and standard petal lacking basal appendages. This morphologically well-defined Neotropical group is also characterized by having twisted

Table 4
Divergence time estimates and ancestral area reconstructions for key nodes. HPD = highest posterior density. Biogeographic areas: A = Northern and Northwestern South America. B = South American dry diagonal; C = Eastern South America; D = Mesoamerica; E = Old World; F = Caribbean. Only the probabilities for the areas estimated as the most likely using the model DEC + j in BioGeoBEARS are shown.

Nodes and major clades	Age estimate and 95% HPD (Mya)	Probabilities for nodal reconstructions
Node 1/*Canavalia*	8.7 11.1–6.7	ACEF: 0.06
Node 2/Clade A	6.2 8.1–4.3	C: 0.40
Node 3/Clade A1	1 1.8–0.5	ACEF: 0.52
Node 4/Clade A2	4.6 6.1–3.2	C: 0.83
Node 5	4 5.4–2.7	C: 0.73
Node 6/Clade B + C	7.6 9.9–5.7	A: 0.16
Node 7/Clade B	4.7 6.7–2.7	ABCDEF: 0.21
Node 8/Clade B1	1.3 2.1–0.7	ABCDEF: 0.66
Node 9/Clade C	3.6 5.1–2.5	A: 0.38
Node 10	2.7 3.4–1.6	A: 0.47
Node 11/Clade C2	2.4 3–1.6	D: 1
Node 12/Clade C3	2.2 3.1–1.4	A: 81
Node 13	1.5 2.2–0.9	A: 0.47 B: 0.27

wing and keel petals, and legumes usually pubescent (Sauer, 1964) (Fig. 4G–I and L). The distinctiveness of the species of this lineage led Piper (1925) to segregate *Wenderothia* as a distinct genus, but our results do not support *Wenderothia* as an independent genus, unless each clade (A, B, and C) was ranked at the generic level.

Although subg. *Catodonia* seems to be a morphologically homogeneous group (Fig. 4A, C and J), it was not recovered as monophyletic in our analyses because of the position of *Canavalia macropleura* and *C. favieri*. *Canavalia macropleura* (subg. *Catodonia*) was recovered as sister to Clade B1 (subg. *Canavalia* + subg. *Maunaloa*). *Canavalia macropleura* and the similar *C. sericophylla* Ducke (not sampled here) differ from subg. *Catodonia* (*sensu* Sauer, 1964) by the free lower calyx teeth and pods with wing-like ribs along the margins (vs. non-expanded ribs along the margins; Piper, 1925). The New Caledonian endemic *Canavalia favieri* (Fig. 4B) was recovered in Clade A1. This species was formerly considered as related to subgenera *Canavalia* or *Maunaloa* because of its calyx morphology (Nielsen, 1993), but it has legumes and ellipsoid seeds similar to those of *C. bonariensis* and *C. nitida* (both in Clade A1), and the lower calyx teeth are partially fused as in other species of subg. *Catodonia*. The only other species from subg. *Catodonia* that was not sampled here is *Canavalia mandibulata*. This species is only known from two old collections from Colombia. The presence of lower calyx teeth partially fused suggests that it should belong to subgenus *Catodonia* (Clade A).

Sauer (1964) distinguished subg. *Maunaloa* from subg. *Canavalia* by the shape of the upper edge of the calyx, the length of the lowest calyx tooth, and the relative size between the keel petals and the standard/wing petals (Sauer, 1964) (Fig. 4D–F and K). However, these characters sometimes overlap, mainly in some species of the Pacific Islands such as *C. megalantha*, *C. sericea*, and *C. vitiensis* (subg. *Canavalia*). Despite these three species seeming to be morphologically more related to the Hawaiian species of subg. *Maunaloa*, Sauer (1964) treated them within subg. *Canavalia*, apparently based more on the geographical distribution than on morphology. Vatanparast et al. (2011) using six chloroplast markers (ca. 6400 bp) found evidence for a paraphyletic subg. *Canavalia* with a monophyletic subg. *Maunaloa* embedded within it whereas Queiroz et al. (2015) found subgenera *Canavalia* and *Maunaloa* reciprocally monophyletic, in accordance with Sauer's (1964) classification. The low interspecific divergence, probably reflects rapid speciation (Vatanparast, 2010; Vatanparast et al., 2011), and the consequent lack of resolution and support within Clade B1 did not allow us to assess the phylogenetic relationships within/between subgenera *Canavalia* and *Maunaloa*. Most unsampled species of *Canavalia* are morphologically linked to Clade B1 and are from the Old World. Because relationships within this clade are mostly unresolved, the inclusion of those unsampled taxa should not change the main conclusion that diversity within this clade resulted from rapid and recent speciation.

4.2. Biogeography

4.2.1. The origin of Canavalia

According to our analyses, *Canavalia* first evolved between 15.8 Mya (stem age) and 8.7 Mya (crown age) during the Miocene. This is much more recent than the Cretaceous origin hypothesized by Sauer (1964). The fossils on which Sauer's (1964) hypotheses relied most likely do not belong to *Canavalia*, and the young age estimate of the genus strengthens this assertion. These ages are older than reported by Li et al. (2013) (stem node = 5 Mya), but that study focused on the diversification of the Phaseoloids and included only a few samples from Diocleae.

4.2.2. South American wet forests

Although the ancestral area of *Canavalia* was reconstructed as equivocal, most of the first diverging species of the major clades (A, B and C) occur in wet forests of South America, leading us to hypothesize that *Canavalia* may have occurred in the wet forests of South America during the Middle Miocene when the global climate was warmer and more humid, and the two large areas of forest (Amazon and Atlantic forests) would have been continuous (Morley, 2000; Zachos et al., 2001). The gradual decrease in temperature and humidity during the Upper Miocene (Zachos et al., 2001) would have resulted in the expansion of the dry forests and savannas and the contraction of wet forests, leading to the segregation of the Amazon and Atlantic forests.

In Clade B, the wet forest ancestors gave origin to Old World lineages through transoceanic dispersal (see Section 4.2.3) and in Clade C they experienced biome evolution to seasonally dry climatic regimes (see Section 4.2.4). Changes in the Northern and Northwestern South American environment after the Andes uplift seem to have played a key role in the origin and diversification of both Clades B and C. The diversification of these lineages coincides with the disappearance of the mega-wetlands of the Pebas system [a huge (>1 million km^2) system of long-lived lakes and wetlands that occupied most of Western Amazonia between ca. 24 and 10 Mya] and the establishment of terrestrial conditions in Western Amazonia about 7 Mya (Antonelli et al., 2009; Hoorn et al., 2010). Similar trends of post-Pebas diversification were also mentioned for other groups of plants, such as *Astrocaryum* G. Mey. (Roncal et al., 2013), *Guatteria* Ruiz & Pav. (Erkens et al., 2007), *Guarea* F. Allam. ex L., *Trichilia* P. Browne (Koenen et al., 2015), and *Inga* Mill. (Richardson et al., 2001).

Species in Clade A are wet forest specialists. Contrary to what was observed in the sea-drift distributed Clade A1 (see Section 4.2.4), the species of Clade A2 currently have a disjunct distri-

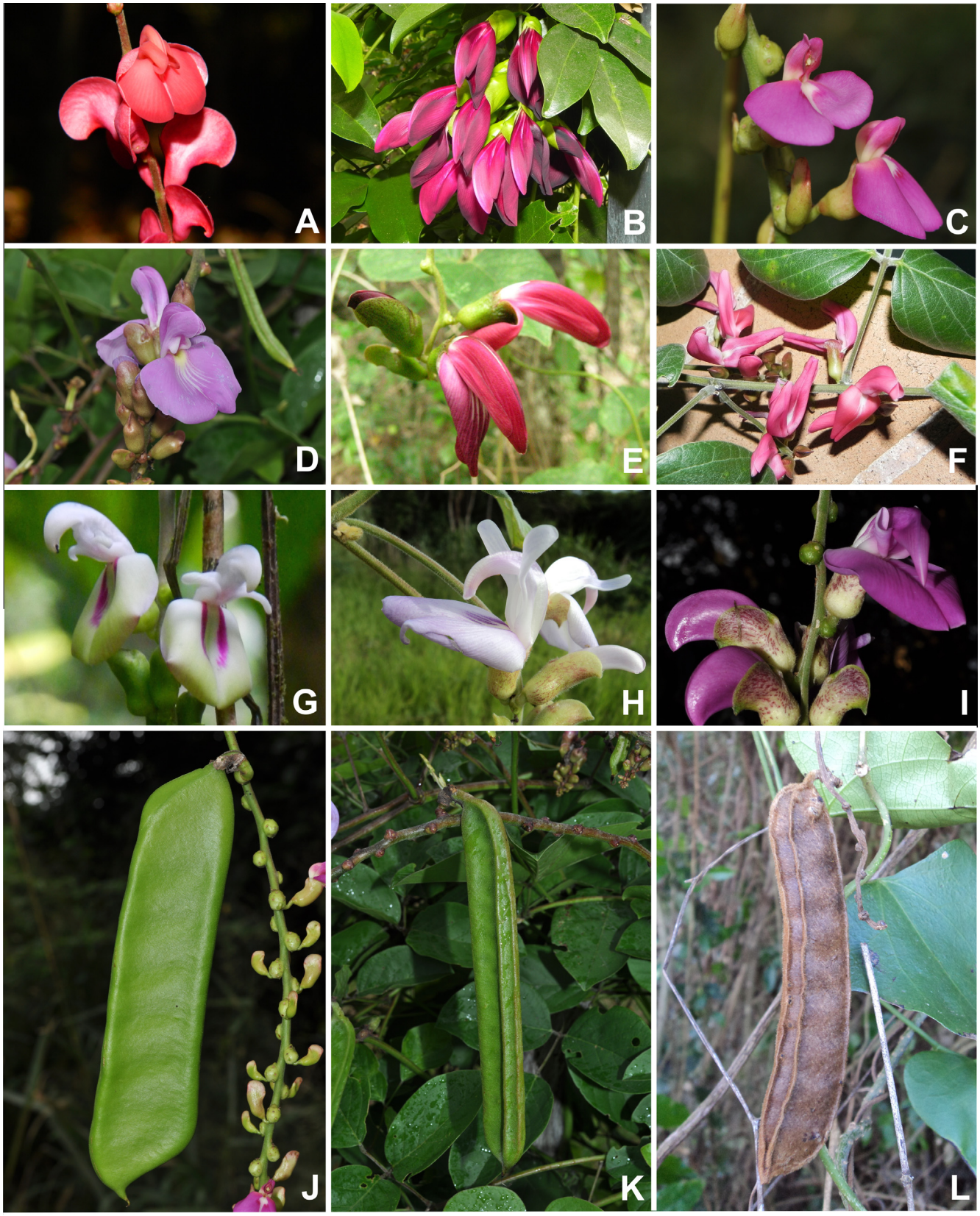

Fig. 4. Flower and fruit diversity in *Canavalia.* A–I: Flower diversity. A: *C. cassidea*; B: *C. favieri*; C: *C. parviflora* (Clade A – Catodonia). D: *C. brasiliensis*; E: *C. hawaiiensis*; F: *C. sericea* (Clade B1 – *Canavalia*/Maunaloa). G: *C. bicarinata*; H: *C. mattogrossensis*; I: *C. villosa* (Clade C – Wenderothia). J–L: Fruit diversity: J: *C. parviflora* (Clade A – Catodonia); K: *C. brasiliensis* (Clade B – *Canavalia*); L: *C. picta* (Clade C – Wenderothia). Photos A, C–D, H–L by C. Snak; B, F by B. Henry; E by K. Kimball; and G by J. Sinasac.

bution between the South American Atlantic and Amazon forests, two large areas of wet forests separated by a diagonal of open and dry vegetation comprising the Caatinga, the Cerrado, and the Chaco (Prado and Gibbs, 1993; Pennington et al., 2004). The Amazonian species of the Clade A1 constitute a monophyletic lineage in Southwestern Amazonia limits in Bolivia and Peru, mainly in the Beni, Madeira, and Madre de Dios River basins. This Amazonian clade is embedded within a paraphyletic grade of Atlantic species distributed in coastal Brazil between 08°S and 20°S (Sauer, 1964; Aymard and Cuello, 1991; Queiroz and Snak, 2015). This Amazonian lineage possibly arose from an Atlantic ancestor that reached Amazonia through a dispersal route linking Southwestern Amazonia and Southeastern Atlantic forest which was inferred to have occurred through gallery forests along the southern margin of the modern Cerrado during the Miocene (Batalha-Filho et al., 2013).

4.2.3. Recent occupancy of areas with seasonally dry climate

The Clade C (subg. *Wenderothia*) probably originated in wet forests from Northern and Northwestern South America between 9.9 Mya and 2.5 Mya (upper and lower 95% HPD limits of stem and crown ages). No species of this clade has buoyant seeds (Sauer, 1964) which explains its restriction to the Neotropics. Two opposite dispersal movements northward and southward followed one event of biome shift toward seasonally dry vegetation (node 9 in Fig. 3).

Northward dispersal to Mesoamerica (mostly Mexico) possibly occurred overland between 3.8 Mya and 1.6 Mya (upper and lower limits of the 95% HPD of the stem and crown node 11), with diversification during the Pleistocene. Such recent speciation in Central American and Mexican dry forests was demonstrated for other dry forest specialists including *Coursetia* DC., *Nissolia* Jacq., and *Ruprechtia* C.A. Mey. (Pennington et al., 2004), contrasting with the well documented older diversification estimation in South American dry forest lineages (Lavin, 2006; Pennington et al., 2009, 2010; Duno de Stefano et al., 2010; Queiroz and Lavin, 2011; Särkinen et al., 2012), suggesting that diversification in Central American dry forests probably followed immigration from South America after the closure of the Isthmus of Panama (see also Section 4.2.4).

The southward migration mostly occurred between 3.8 Mya and 1.4 Mya (upper and lower 95% HPD limits of the stem and crown node 12) and resulted in the occupancy of areas subject to seasonal climate mostly in Central Brazil. This area consists mostly of seasonally dry, fire prone savanna vegetation on dystrophic soils of the Cerrado domain. This relatively young age is in line with the finding that most lineages of the Cerrado started to diversify at 4 Mya or less, coinciding with the establishment of fire regimes in that region (Simon et al., 2009). The specialization of the species of Clade C to dystrophic soil conditions is underscored by species that migrated to the Atlantic Forest and back to the Amazon region (*C. picta* and *C. albiflora*, respectively) but are restricted to drier and open areas.

4.2.4. Intercontinental disjunctions and the age of the Isthmus of Panama

The finding that *Canavalia* arose about 15.8 Mya allowed us to reject the idea that the genus reached its pantropical distribution by vicariance of continental land masses, as suggested by Sauer (1964), or through an early Tertiary biotic interchange between North America and Eurasia (Lavin and Luckow, 1993).

Our data strongly support recent transoceanic dispersal and we argue that a trans-Pacific dispersal is the scenario that best explains the current distribution patterns in the genus. This view is supported by at least two independent dispersal events from the New World to the Old World (Fig. 3; Appendix G); each Old World lineage is related with a clade with buoyant seeds (see ancestral reconstruction of seed buoyancy in Appendix H); all intercontinental disjunctions have a mean age less than 6.2 Ma (Clade A) or 1.3 Ma (Clade B1; Fig. 3; Appendix E); finally, sea currents had a westward direction at that time (Renner, 2004). The failure to include all Old World species of subgenera *Canavalia* and *Maunaloa* (clade B1; see Section 4.1) should not change the main result found here of a recent Trans-Pacific dispersal because the unsampled species are distributed in the same broad area along Pacific and Indian oceans and relationships within clade B1 are rather unresolved because of low molecular variation among their species.

Intercontinental disjunctions were observed in clade A1 (*C. bonariensis* in South America, Eastern South Africa, Madagascar, and Reunión, *C. nitida* in the Caribbean, and *C. favieri* in New Caledonia) and in clade B1 (see Fig. 3). Several species of these two clades have buoyant and impermeable seeds (Sauer, 1964; Beyra-Matos et al., 2004) which may have enabled these lineages to disperse via sea currents.

No New World Clade A species with a close relationship to Old World lineages occur along the Pacific coast. Instead, they are distributed in the Caribbean or along the Atlantic coast of Southern South America. This raises a question about the dispersal route between the Atlantic and Pacific oceans. The closure of the Isthmus of Panama has been dated at around 3.5 Mya by several studies (Coates et al., 1992, 2003; Coates and Stallard, 2013; Molnar, 2008; Hoorn et al., 2010). This has been recently challenged by studies using stratigraphic, molecular and fossil data, which suggest extensive terrestrial landscapes were already formed in the region of the Isthmus of Panama by 20 Mya (Bacon et al., 2015; Hoorn and Flantua, 2015; Montes et al., 2015), although authors assumed that these land masses were not necessarily fully connected. Our data suggest that a seaway linking the Atlantic and Pacific oceans was still open ca. 8 Mya (the lower limits of the 95% HPD distribution), even if the emergence and closure of the Isthmus of Panama was close to completion. Easterly trade winds into the open Central American Seaway drove surface currents westward from the Caribbean to the Pacific (Sepulchre et al., 2014) and could have carried buoyant seeds of the species of Clade A1, resulting in the present biogeographic pattern of this clade.

5. Conclusions

Our results demonstrate that of the four subgenera of *Canavalia* proposed by Sauer (1964), only subg. *Wenderothia* is monophyletic. To render subg. *Catodonia* monophyletic, it is necessary to include the New Caledonian *Canavalia favieri* and exclude the Amazonian *C. macropleura*. Our results are not conclusive about relationships between subgenera *Canavalia* and *Maunaloa*.

The divergence time estimate analyses indicated that *Canavalia* arose during the Miocene and that most of its extant species diverged rapidly during the Pleistocene. The divergence of the major clades is chronologically consistent with different climatic (cooler and drier vs. warmer and humid periods) and geological (the Andes uplift, the formation of the Pebas and post-Pebas systems, and the closure of the Isthmus of Panama) events, and the change in the direction of ocean currents. The dispersal of the subg. *Catodonia* lineage from the New World to the Old World through the Central American Seaway supports the hypothesis that a seaway linking Atlantic and Pacific oceans was open at ca. 8 Mya. The genus *Canavalia* achieved its present day pantropical distribution through recent transoceanic dispersal, but more variable molecular markers for subgenera *Canavalia* and *Maunaloa* and the inclusion of the remaining species are needed to estimate the number of times buoyant seeds evolved and the number of disper-

sal events between the New World and Old World. It is evident that the phylogeny and geography of *Canavalia* cannot be reconciled by a single historical event, but rather by a complex set of processes.

Acknowledgments

We thank the curators of the herbaria visited for access to specimens; the herbaria B, BM, CEN, CEPEC, F, G, GH, HUEFS, K, M, MBM, MO, NY, P, and US for the donation of DNA samples; Ariane Barbosa, Domingos Cardoso, Leonardo Borges, Maria Cristina Roberts, and Wallace São-Mateus for leaf material of *Canavalia*; Centro Internacional de Agricultura Tropical (CIAT, Cali, Colombia) for providing seeds of the species of *Canavalia* from the forage germplasm collection; Alan Graham and Patrick Herendeen for the help with fossils; Alexandre Zuntini for the help with analysis in BioGeoBEARS; Benoît Henry for the photos of *C. favieri* and *C. sericea*; Jenn Sinasac for the photos of *C. bicarinata*; Kanoa Kimball for the photos of *C. hawaiiensis*; Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) for issuing collecting permits. Fieldwork and DNA sequencing were sponsored by the Programa de Pesquisa em Biodiversidade do Semiárido (PPBIO), Sistema Nacional de Pesquisa em Biodiversidade (SISBIOTA CNPq 563084/2010-3/FAPESB PES0053/2011), PRONEX (FAPESB PNX0014/2009), PROTAX (CNPq 562354/2010-7), and REFLORA (CNPq 563546/2010-7/FAPESB PES0054/2011). This study was also supported by the Environment Research and Technology Development Fund (S-9) of the Ministry of the Environment, Japan. C. Snak also thanks the Missouri Botanical Garden for the 2014 Elizabeth E. Bascom Fellowship. This study is part of the first author's PhD thesis developed in the Programa de Pós-Graduação em Botânica (UEFS) and supported by PhD fellowship grants from CAPES, and SWE CAPES at the Royal Botanic Gardens, Kew.

Appendices A–H. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2016.02.001.

References

Antonelli, A., Nylander, J.A.A., Persson, C., Sanmartín, I., 2009. Tracing the impact of the Andean uplift on Neotropical plant evolution. Proc. Natl. Acad. Sci. U.S.A. 106, 9749–9754.

Aymard, G.C., Cuello, N.A., 1991. Catalogo y adiciones a las species neotropicales del genero *Canavalia* (Leguminosae–Papilionoideae–Phaseoleae–Diocleinae). In: Vargas, R.E., León, A., Escobar, A. (Eds.), Compilación de los trabajos presentados en el primer Seminario-Taller sobre *Canavalia* ensiformis. Maracay, Venezuela, pp. 45–64.

Bacon, C.D., Silvestro, D., Jaramillo, C., Smith, B.T., Chakrabarty, P., Antonelli, A., 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. Proc. Natl. Acad. Sci. U.S.A. 112, 6110–6115.

Baldwin, B.G., Markos, S., 1998. Phylogenetic utility of the external transcribed spacers (ETS) of 18S–26S rDNA: congruence of ETS and ITS trees of *Calycadenia* (Compositae). Mol. Phylogenet. Evol. 10, 449–463.

Batalha-Filho, H., Fjeldsa, J., Fabre, P.H., Muyaki, C.Y., 2013. Connections between the Atlantic and Amazonian Forest avifaunas represent distinct historical events. J. Ornithol. 154, 41–50.

Bentham, G., 1837. Commentationes de Leguminosarum Generibus. Zollinger, Vienna.

Berry, E.W., 1916. The lower Eocene floras of southeastern North America. U.S. Geol. Surv. Prof. Pap. 91, 1–353.

Berry, E.W., 1925. The Tertiary flora of the island of Trinidad. B.W.I. Johns Hopkins Univ. Stud. Geol. 6, 71–161.

Berry, E.W., 1930. Revision of the lower Eocene Wilcox flora of the southeastern states. U.S. Geol. Surv. Prof. Pap. 156, 1–144.

Beyra-Matos, A., Reyes, G., Hernández, L., Herrera, P., 2004. Revisión taxonómica del género *Canavalia* DC. (Leguminosae–Papilionoideae) em Cuba. Rev. Acad. Colomb. Cienc. 28, 157–175.

Cardoso, D., Queiroz, L.P., Pennington, R.T., Lima, H.C., Fonty, E., Wojciechowski, M. F., Lavin, M., 2012. Revisiting the phylogeny of papilionoid legumes: new insights from comprehensively sampled early-branching lineages. Am. J. Bot. 99, 1991–2013.

Cardoso, D., Pennington, R.T., Queiroz, L.P., Boatwright, J.S., Van Wyk, B.-E., Wojciechowski, M.F., Lavin, M., 2013. Reconstructing the deep-branching relationships of the papilionoid legumes. S. Afr. J. Bot. 89, 58–75.

Cardoso, D., São-Mateus, W.M.B., da Cruz, D.T., Zartman, C.E., Komura, D.L., Kite, G., Prenner, G., Wieringa, J.J., Clark, A., Lewis, G.P., Pennington, R.T., Queiroz, L.P., 2015. Filling in the gaps of the papilionoid legume phylogeny: the enigmatic Amazonian genus *Petaladeniu*m is a new branch of the early-diverging Amburaneae clade. Mol. Phylogenet. Evol. 84, 112–124.

Coates, A.G., Jackson, J.B.C., Collins, L.S., Cronin, T.M., Dowsett, H.J., Bybell, L.M., Jung, P., Obando, J.A., 1992. Closure of the Isthmus of Panama: the near-shore marine record of Costa Rica and western Panama. Geol. Soc. Am. Bull. 104, 814–828.

Coates, A.G., Aubry, M.P., Berggren, W.A., Collins, L.S., Kunk, M., 2003. Early Neogene history of the Central American arc from Bocas del Toro, western Panama. Geol. Soc. Am. Bull. 115, 271–287.

Coates, A.G., Stallard, R.F., 2013. How old is the Isthmus of Panama. Bull. Mar. Sci. 89, 801–814.

Cohen, K.M., Finney, S.C., Gibbard, P.L., Fan, J., 2013. The ICS international chronostratigraphic chart. Episodes 36, 199–204 (v2015/01).

Côrtes, A.L.A., Rapini, A., Daniel, T.F., 2015. The *Tetramerium* lineage (Acanthaceae: Justicieae) does not support the Pleistocene Arc hypothesis for South American seasonally dry forests. Am. J. Bot. 102, 992–1007.

Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation method for small quantities of fresh tissues. Phytochem. Bull. 19, 11–15.

Doyle, J.J., Doyle, J.L., 1993. Chloroplast DNA phylogeny of the papilionoid legume tribe Phaseoleae. Syst. Bot. 18, 309–327.

Drummond, C.S., 2008. Diversification of *Lupinus* (Leguminosae) in the western New World: derived evolution of perennial life history and colonization of montane habitats. Mol. Phylogenet. Evol. 48, 408–421.

Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., Wilson, A., 2012a. Geneious Version 6.1.6. <http://www.geneious.com>.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012b. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

Duno de Stefano, R., Fernández-Concha, G.C., Can-Itza, L.L., Lavin, M., 2010. The morphological and phylogenetic distinctions of *Coursetia greenmanii* Leguminosae): taxonomic and ecological implications. Syst. Bot. 35, 289–295.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl. Acids Res. 32, 1792–1797.

Ekanayake, S., Skog, K., Asp, N.G., 2007. Canavanine content in sword beans (*Canavalia gladiata*): analysis and effect of processing. Food Chem. Toxicol. 45, 797–803.

Erixon, P., Svennblad, B., Britton, T., Oxelman, B., 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst. Biol. 52, 665–673.

Erkens, R.H.J., Chatrou, L.W., Maas, J.W., Van der Niet, T., Savolainen, V., 2007. A rapid diversification of rainforest trees (*Guatteria*: Annonaceae) following dispersal from Central into South America. Mol. Phylogenet. Evol. 44, 399–411.

Farris, J., Kallersjo, M., Kluge, A.G., Bult, C., 1995. Testing significance of incongruence. Cladistics 10, 315–319.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Fitch, W.M., 1971. Towards defining the course of evolution: minimum change for a specific tree topology. Syst. Zool. 20, 406–416.

Hedges, S.B., 1992. The number of replications needed for accurate estimation of the bootstrap *p* value in phylogenetic studies. Mol. Biol. Evol. 9, 366–369.

Herendeen, P.S., 1992. The fossil history of the Leguminosae from the Eocene of Southeastern North America. In: Herendeen, P.S., Dilcher, D.L. (Eds.), Advances in Legume Systematics, Part 4, the Fossil Record. Royal Botanic Gardens, Kew, pp. 85–160.

Hoorn, C., Wesselingh, F.P., ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., Sanmartin, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., Jaramillo, C., Riff, D., Negri, F.R., Hooghiemstra, H., Lundberg, J., Stadler, T., Sarkinen, T., Antonelli, A., 2010. Amazonia through time: andean uplift, climate change, landscape evolution, and biodiversity. Science 330, 927–931.

Hoorn, C., Flantua, S., 2015. An early start for the Panama land bridge. Science 348, 186–187.

Hu, J., Lavin, M., Wojciechowski, M.F., Sanderson, M.J., 2000. Phylogenetic systematics of the tribe Millettieae (Leguminosae) based on chloroplast trnK/matK sequences and its implications for evolutionary patterns in the Papilionoideae. Am. J. Bot. 87, 418–430.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P., 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294, 2310–2314.

Hughes, C.E., Nyffeler, R., Linder, H.P., 2015. Evolutionary plant radiations: where, when, why and how? New Phytol. 207, 249–253.

Kajita, T., Ohashi, H., Tateishi, Y., Bailey, C.D., Doyle, J.J., 2001. RbcL and legume phylogeny, with particular reference to Phaseoleae, Millettieae, and allies. Syst. Bot. 26, 515–536.

Koenen, E.J.M., Clarkson, J.J., Pennington, T.D., Chatrou, L.W., 2015. Recently evolved diversity and convergent radiations of rainforest mahoganies (Meliaceae) shed new light on the origins of rainforest hyperdiversity. New Phytol. 207, 327–339.

Kollipara, K.P., Singh, R.J., Hymowitz, T., 1997. Phylogenetic and genomic relationships in the genus *Glycine* Willd. based on sequences from the ITS region of nuclear rDNA. Genome 40, 57–68.

Kress, W.J., Prince, L.M., Williams, K.J., 2002. The phylogeny and a new classification of the gingers (Zingiberaceae): evidence from molecular data. Am. J. Bot. 89, 1682–1696.
Landis, M.J., Matzke, N.J., Moore, B.R., Huelsenbeck, J.P., 2013. Bayesian analysis of biogeography when the number of areas is large. Syst. Biol. 62, 789–804.
Lavin, M., Luckow, M., 1993. Origins and relationships of tropical North America in the context of the boreotropics hypothesis. Am. J. Bot. 80, 1–14.
Lavin, M., Herendeen, P., Wojciechowski, M., 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. Syst. Biol. 54, 575–594.
Lavin, M., 2006. Floristic and geographical stability of discontinuous seasonally dry tropical forests explain patterns of plant phylogeny and endemism. In: Pennington, R.T., Lewis, G.P., Ratter, J.A. (Eds.), Neotropical Savannas and Seasonally Dry Forests: Plant Diversity, Biogeography and Conservation. CRC Press, Taylor & Francis Group, Boca Raton, London, New York, pp. 433–447.
Li, H., Wang, W., Lin, L., Zhu, X., Li, J., Zhu, X., Chen, Z., 2013. Diversification of the phaseoloid legumes: effects of climate change, range expansion and habit shift. Front. Plant Sci. 4, 386.
Lohmann, L.G., Bell, C.D., Calió, M.F., Winkworth, R.C., 2013. Pattern and timing of biogeographical history in the Neotropical tribe Bignonieae (Bignoniaceae). Bot. J. Linn. Soc. 171, 154–170.
LPWG [Legume Phylogeny Working Group], 2013. Legume phylogeny and classification in the 21st century: progress, prospects and lessons for other species-rich clades. Taxon 62, 217–248.
Maddison, W.P., Maddison, D.R., 2015. Mesquite: a Modular System for Evolutionary Analysis. Version 3.02. <http://mesquiteproject.org>.
Matzke, N.J., 2013. BioGeoBEARS: Biogeography with Bayesian (and likelihood) Evolutionary Analysis in R Scripts. CRAN: The Comprehensive R Archive Network.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), November 2010, New Orleans, pp. 1–8.
Molnar, P., 2008. Closing of the Central American Seaway and the ice age: a critical review. Paleoceanography 23, PA2201.
Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J.C., Valencia, V., Ayala, C., Pérez-Angel, L.C., Rodriguez-Parra, L.A., Ramirez, V., Niño, H., 2015. Middle miocene closure of the Central American seaway. Science 384, 226–229.
Morley, R.J., 2000. Origin and Evolution of Tropical Rain Forests. Wiley, Chichester, UK.
Müller, K., 2005a. SeqState – primer design and sequence statistics for phylogenetic DNA data sets. Appl. Bioinform. 4, 65–69.
Müller, K.F., 2005b. The efficiency of different search strategies in estimating parsimony jackknife, bootstrap, and Bremer support. BMC Evol. Biol. 5, 58.
NAS–National Academy of Sciences, 1979. Tropical Legumes: Resources for the Future. National Academy of Sciences, Washington.
Nielsen, I., 1993. A new species of *Canavalia* (Leguminosae–Papilionoideae) from New Caledonia. Bull. Mus. Natl. Hist. Nat. B 14, 347–350.
Nylander, J.A.A., 2004. MrModeltest v. 2. Program Distributed by the Author. Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden.
Paithankar, K.R., Prasad, K.S.N., 1991. Precipitation of DNA by polyethylene glycol and ethanol. Nucl. Acids Res. 19, 1346.
Perez, A.P.F., Silva, M.J., Queiroz, L.P., Lewis, G.P., Simões, A.O., Azevedo-Tozzi, A.M. G., Sarkinen, T.E., Souza, A.P., 2013. Phylogeny and biogeography of the genus *Zornia* (Leguminosae: Papilionoideae: Dalbergieae). Taxon 62, 723–733.
Pennington, R.T., Lavin, M., Prado, D.E., Pendry, C.A., Pell, S.K., Butterworth, C.A., 2004. Historical climate change and speciation: neotropical seasonally dry forest plants show patterns of both Tertiary and Quaternary diversification. Philos. Trans. R. Soc. Lond. B 359, 515–537.
Pennington, R.T., Lavin, M., Oliveira-Filho, A., 2009. Woody plant diversity, evolution and ecology in the tropics: perspectives from seasonally dry tropical forests. Ann. Rev. Ecol. Evol. Syst. 40, 437–457.
Pennington, R.T., Lavin, M., Särkinen, T., Lewis, G.P., Klitgaard, B.B., Hughes, C.E., 2010. Contrasting plant diversification histories within the Andean biodiversity hotspot. Proc. Natl. Acad. Sci. U.S.A. 107, 13783–13787.
Piper, C.V., 1925. The American species of *Canavalia* and *Wenderothia*. Contr. U.S. Nat. Herb. 20, 555–588.
Prado, D.E., Gibbs, P.E., 1993. Patterns of species distributions in the dry seasonal forests of South America. Ann. Mo. Bot. Gard. 80, 902–927.
Queiroz, L.P., Fortunato, R.H., Giulietti, A.M., 2003. Phylogeny of Diocleinae (Papilionoideae: Phaseoleae) based on morphological characters. In: Klitgaard, B.B., Bruneau, A. (Eds.), Advances in Legume Systematics, Part 10, Higher Level Systematics. Royal Botanic Gardens, Kew, pp. 303–324.
Queiroz, L.P., Lewis, G.P., Wojciechowski, M.F., 2010. *Tabaroa*, a new genus of Leguminosae tribe Brongniartieae from Brazil. Kew Bull. 65, 189–203.
Queiroz, L.P., Lavin, M., 2011. *Coursetia* (Leguminosae) from eastern Brazil: nuclear ribosomal and chloroplast DNA sequence analysis reveal the monophyly of three Caatinga-inhabiting species. Syst. Bot. 36, 69–79.
Queiroz, L.P., Pastore, J.F.B., Cardoso, D., Snak, C., Lima, A.L.C., Gagnon, E., Vatanparast, M., Holland, A.E., Egan, A.N., 2015. A multilocus phylogenetic analysis reveals the monophyly of a recircumscribed papilionoid legume tribe Diocleae with well-supported generic relationships. Mol. Phylogenet. Evol. 90, 1–19.
Queiroz, L.P., Snak, C., 2015. *Canavalia* in Lista de Espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro. <http://floradobrasil.jbrj.gov.br/jabot/floradobrasil/FB22854> (access: 20 Jul. 2015).
R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL: <https://www.r-project.org/>.
Rambaut, A., 2009. FigTree version 1.3.1: tree figure drawing tool. Computer Program and Documentation Distributed by the Author. <http://tree.bio.ed.ac.uk/software/figtree>.
Rambaut, A., Drummond, A.J., 2013. Tracer version 1.6. Computer Program and Documentation Distributed by the Author. <http://beast.bio.ed.ac.uk/Tracer>.
Ree, R.H., Smith, S.A., 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Syst. Biol. 57, 4–14.
Renner, S., 2004. Plant dispersal across the Tropical Atlantic by wind and sea currents. Int. J. Plant Sci. 165, S23–S33.
Richardson, J.E., Pennington, R.T., Pennington, T.D., Hollingsworth, P.M., 2001. Rapid diversification of a species-rich genus of Neotropical rain forest trees. Science 293, 2242–2245.
Roncal, J., Kahn, F., Millan, B., Couvreur, T.L.P., Pintaud, J.-C., 2013. Cenozoic colonization and diversification patterns of tropical American palms: evidence from *Astrocaryum* (Arecaceae). Bot. J. Linn. Soc. 171, 120–139.
Ronquist, F., 1997. Dispersal–vicariance analysis: a new approach to the quantification of historical biogeography. Syst. Biol. 46, 195–203.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Samarakoon, T., Wang, S.Y., Alford, M.H., 2013. Enhancing PCR amplification of DNA from recalcitrant plant specimens using a trehalose-based additive. Appl. Plant Sci. 1, 1200236.
Särkinen, T., Pennington, R.T., Lavin, M., Simon, M.F., Hughes, C.E., 2012. Evolutionary islands in the Andes: persistence and isolation explain high endemism in Andean dry tropical forests. J. Biogeogr. 39, 884–900.
Sauer, J., 1964. Revision of *Canavalia*. Brittonia 16, 106–181.
Sepulchre, P., Arsouze, T., Donnadieu, Y., Dutay, J.-C., Jaramillo, C., Le Bras, J., Martin, E., Montes, C., Waite, A.J., 2014. Consequences of shoaling of the Central American Seaway determined from modeling Nd isotopes. Paleoceanography 29, 176–189.
Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.
Silva, M.J., Queiroz, L.P., Azevedo-Tozzi, A.M.G., Lewis, G.P., Souza, A.P., 2012. Phylogeny and biogeography of *Lonchocarpus* sensu lato and its allies in the tribe Millettieae (Leguminosae, Papilionoideae). Taxon 61, 93–108.
Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49, 369–381.
Simon, M.F., Grether, R., Queiroz, L.P., Skema, C., Pennington, R.T., Hughes, C.E., 2009. Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. Proc. Natl. Acad. Sci. U.S.A. 106, 20359–20364.
Simon, M.F., Pennington, T., 2012. Evidence for adaptation to fire regimes in the tropical savannas of the Brazilian Cerrado. Int. J. Plant Sci. 173, 711–723.
Snak, C., 2015. Filogenia e Biogeografia de *Canavalia* Adans. (Leguminosae–Papilionoideae) e Taxonomia das Espécies Neotropicais. PhD Thesis. Universidade Estadual de Feira de Santana, Brazil.
Souza, E.R., Lewis, G.P., Forest, F., Schnadelbach, A.S., van den Berg, C., Queiroz, L.P., 2013. Phylogeny of *Calliandra* (Leguminosae: Mimosoideae) based on nuclear and plastid molecular markers. Taxon 62, 1200–1219.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
Sun, Y., Skinner, D.Z., Liang, G.H., Hulbert, S.H., 1994. Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. Theor. Appl. Genet. 89, 26–32.
Swofford, D.L., 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4.0b10. Sinauer, Sunderland, Massachusetts, USA.
Varela, E.S., Lima, J.P.M.S., Galdino, A.S., Pinto, L.S., Bezerra, W.M., Nunes, E.P., Alves, M.A.O., Grangeiro, T.B., 2004. Relationships in subtribe Diocleinae (Leguminosae; Papilionoideae) inferred from Internal Transcribed Spacer sequences from nuclear ribosomal DNA. Phytochemistry 65, 59–69.
Vatanparast, M., 2010. Phylogeography of a pantropical plant with sea-drifted seeds; *Canavalia* rosea (Sw.) DC. (Fabaceae). PhD Thesis. Chiba University, Japan.
Vatanparast, M., Takayana, K., Sousa, M.S., Tateishi, Y., Kajita, T., 2011. Origin of Hawaiian endemic species of *Canavalia* (Fabaceae) from sea-dispersed species revealed by chloroplast and nuclear DNA sequences. J. Jpn. Bot. 86, 15–25.
Wojciechowski, M.F., Lavin, M., Sanderson, M., 2004. A phylogeny of legumes (Leguminosae) based on analysis of the plastid matK gene resolves many well-supported subclades within the family. Am. J. Bot. 91, 1846–1862.
Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292, 686–693.