The LINNEAN SOCIETY of London BOTANICAL Journal of the Linnean Society

Botanical Journal of the Linnean Society, 2016, **181**, 171–198. With 8 figures.

Phylogenetics, delimitation and historical biogeography of the pantropical tree genus *Thespesia* (Malvaceae, Gossypieae)

F. ARECES-BERAZAIN* and J. D. ACKERMAN

Department of Biology, University of Puerto Rico, PO Box 23360, San Juan, PR 00931-3360

Received 31 October 2015; revised 4 February 2016; accepted for publication 23 February 2016

Thespesia consists of 16 species of trees and shrubs from Southeast Asia–Oceania, Africa and America, the most well known being *T. populnea*, a small tree of tropical coastal areas around the world. Phylogenetic relationships in the genus and among its allies in tribe Gossypieae were inferred using three plastid and two nuclear regions to ascertain its generic delimitation and explore its biogeographical history. Maximum-likelihood and Bayesian analyses confirmed that *Thespesia* is not monophyletic and, based on these results, *Azanza* is reinstated to accommodate the two species previously placed in *Thespesia* section *Lampas*. Dating analyses and ancestral range estimation indicated that *Thespesia s.s.* most likely originated in Southeast Asia–Oceania *c.* 30 Mya, but extant species did not begin to differentiate until the late Miocene. Two dispersal events, one into Africa *c.* 11 Mya and another into America (Antilles) *c.* 9 Mya, gave rise to the African and the Greater Antillean endemics, respectively. The two most widespread hydrochorous species, *T. populnea* and *T. populneoides*, originated in Southeast Asia–Oceania from where they spread to other parts of the world. Our analysis also indicated a much earlier origin than previously reported for the eumalvoid clade and its tribes Gossypieae, Malveae and Hibisceae suggesting that vicariance might have had an important role early in the history of these groups. *Botanical Journal of the Linnean Society*, 2016, **181**, 171–198

ADDITIONAL KEYWORDS: *Azanza* – *Cephalohibiscus* – eumalvoids – fossilized birth death model – long-distance dispersal – Malvoideae – relaxed molecular clock – *Thepparatia*.

INTRODUCTION

With their juxtaposition between continents and a complex geological history, the Caribbean islands have long been viewed as an ideal setting for studying the processes that generate biodiversity. The islands are renowned for their exceptional plant diversity and endemism (at the generic and species levels), which are the product of two contrasting factors: the close proximity to continents and the high degree of isolation.

The role of neighbouring continents as a primary source of colonizing biota has long been recognized by authors who have compiled plant distribution data for the region (Alain, 1958; Howard, 1973; Borhidi, 1991; Acevedo-Rodríguez & Strong, 2008). More recently, molecular phylogenetic studies have shown that the closest extant relatives of many Caribbean insular plant taxa are indeed found in South, Central and North America (see reviews by Santiago-Valentin & Olmstead, 2004; Maunder *et al.*, 2011; Oleas *et al.*, 2013). A few studies, however, have also revealed unexpected biogeographical links with the Old World, notably tropical Africa and the Pacific (Motley, Wurdack & Delprete, 2005; Mort *et al.*, 2008; Andrus *et al.*, 2009; Graham, 2010; Namoff *et al.*, 2010), indicating that our understanding of the processes that have shaped the Caribbean insular flora is still far from adequate.

This study focuses on *Thespesia* Sol. ex Corrêa, a pantropical genus of trees and shrubs consisting of 16–17 species (Fryxell, 1979; Verdcourt & Mwachala, 2009). Only four *Thespesia* spp. occur in the New World, three of which are endemic to the Greater Antilles: *T. cubensis* (Britton & P.Wilson) J.B.Hutch. (Fig. 1A) in Cuba, *T. beatensis* (Urb.) Fryxell (Fig. 1B) in Beata Island (off the southern coast of the Barahona peninsula, Hispaniola) and *T. grandiflora*

*Corresponding author. E-mail: fabiareces@yahoo.com

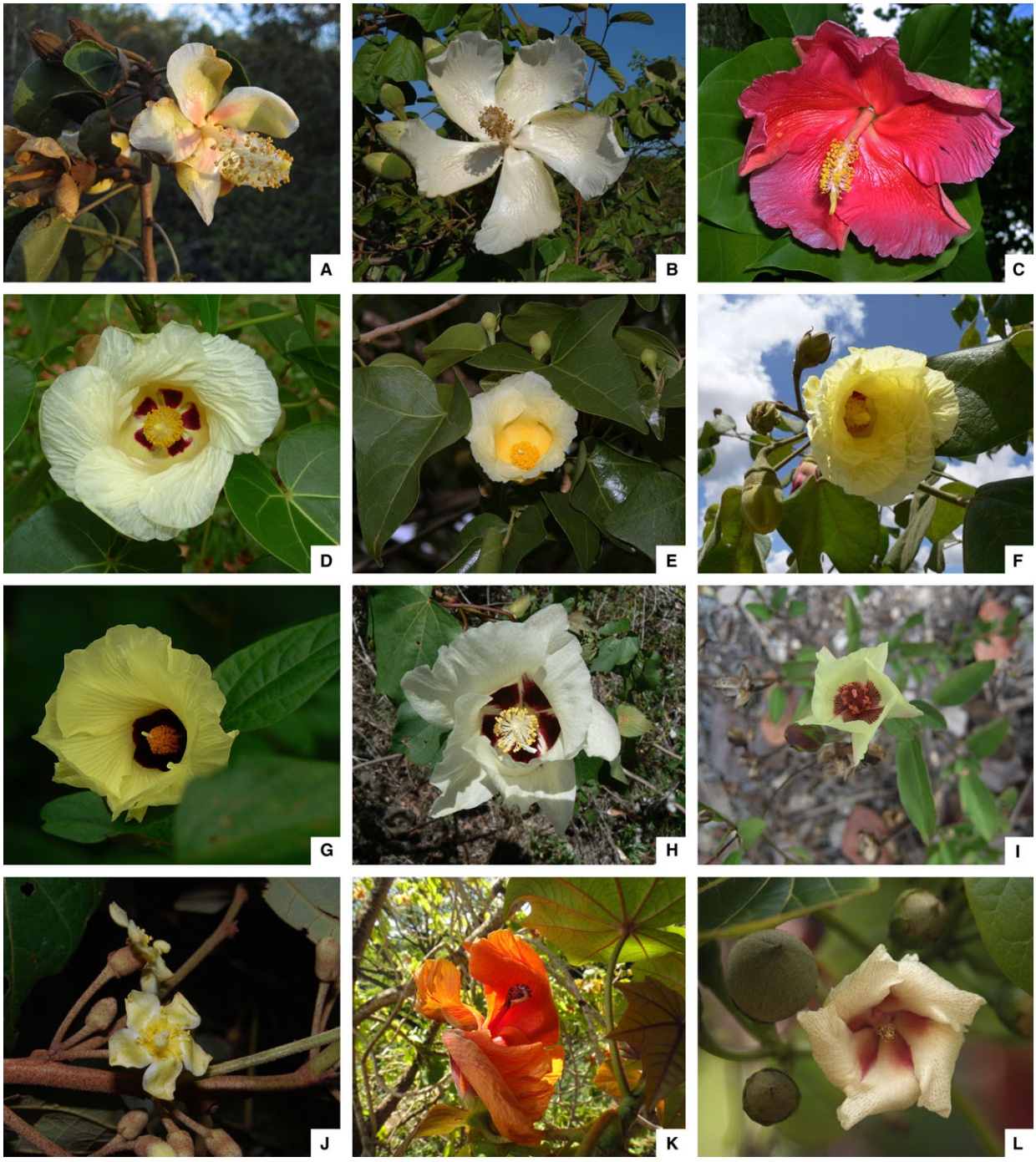

Figure 1. Representative members of *Thespesia* and tribe Gossypieae: A, *T. cubensis* (Cuba); B, *T. beatensis* (Beata Island, Dominican Republic); C, *T. grandiflora* (Puerto Rico); D, *T. populnea* (Puerto Rico); E, *T. acutiloba* (South Africa); F, *T. garckeana* (Malawi); G, *T. lampas* (India); H, *Gossypium hirsutum* (Puerto Rico); I, *Cienfuegosia heterophylla* (Cuba); J, *Hampea appendiculata* (Costa Rica); K, *Kokia kauaiensis* (Hawaii); L, *Lebronnecia kokioides* (Hawaii). Photographs A and I by P. A. González Gutiérrez, B–D and H by F. Areces-Berazain, E by C. Gibbon, F by G. Baumann, G by D. Valke, J by A. Monro and K–L by W. Recart.

DC. (Fig. 1C) in Puerto Rico. The fourth, *T. populnea* (L.) Sol. ex Corrêa (Fig. 1D), is widely distributed in the Caribbean and throughout the Tropics (Fig. 2). The remaining species are concentrated in East Africa and Southeast Asia and many of them also have restricted, isolated ranges. Five species are

endemic to Africa, including one in Madagascar; four are endemic to New Guinea and one is endemic to Australia (Figs 2, 3). Two other species are widespread in Southeast Asia and the coasts of the Indian Ocean, respectively.

Members of *Thespesia* are valued for their multiple uses, ranging from medicinal and food to ornamental, timber and fibre, and *T. populnea* is an important plant in the culture of many countries of the Indo-Pacific; for example, in Polynesia, it is considered a sacred tree (Bates, 1990; Friday & Okano, 2006). The species has recently received extensive attention because of its many medicinal properties (Vasudevan & Parle, 2007; Shrivastav *et al.*, 2009; Parthasarathy, Ilavarasan & Karrunakaran, 2009; Dhanarasu *et al.*, 2010; Selva Kumar *et al.*, 2012; Pratap Chandran *et al.*, 2014). *Thespesia grandiflora*, the state flower of Puerto Rico, is widely planted as an ornamental tree. Its wood is hard and durable and has been used for furniture and craft items (Little & Wadsworth, 1964; Francis, 1999). *Thespesia garckeana* F.Hoffm. (Fig. 1F) is appreciated for its sweet fruits, which are eaten in several African countries (Mojeremane & Tshwenyane, 2004). *Thespesia lampas* (Cav.) Dalzell (Fig. 1G) is widely used in Asia as an ornamental and medicinal plant (Kosalge & Fursule, 2009; Chumbhale & Upasani, 2012a,b).

Thespesia belongs to tribe Gossypieae (Malvaceae, Malvoideae), a well-defined, natural group including the commercially important species of cotton (*Gossypium* L.) (Fryxell, 1979; Seelanan, Schnabel & Wendel, 1997). Members of Gossypieae are distinguished by embryo morphology (the cotyledons are distinctly folded enclosing the epicotyl and hypocotyl) and the presence of gossypol glands (Fryxell, 1968, 1979), which are lysigenous cavities that contain sesquiterpenoid aldehydes that are presumably involved in defence against herbivores (Bell *et al.*, 1975). Other features shared by the genera in the tribe are the presence of an undivided style (except in *Gossypioides* Skovsted and some species of *Cienfuegosia* Cav.) and capsular fruits with three to five carpels (Fryxell, 1979).

Gossypieae currently include eight genera and *c.* 126 species native to tropical and subtropical regions all over the world (Fryxell, 1979). The largest genus is *Gossypium* (Fig. 1H), which includes some

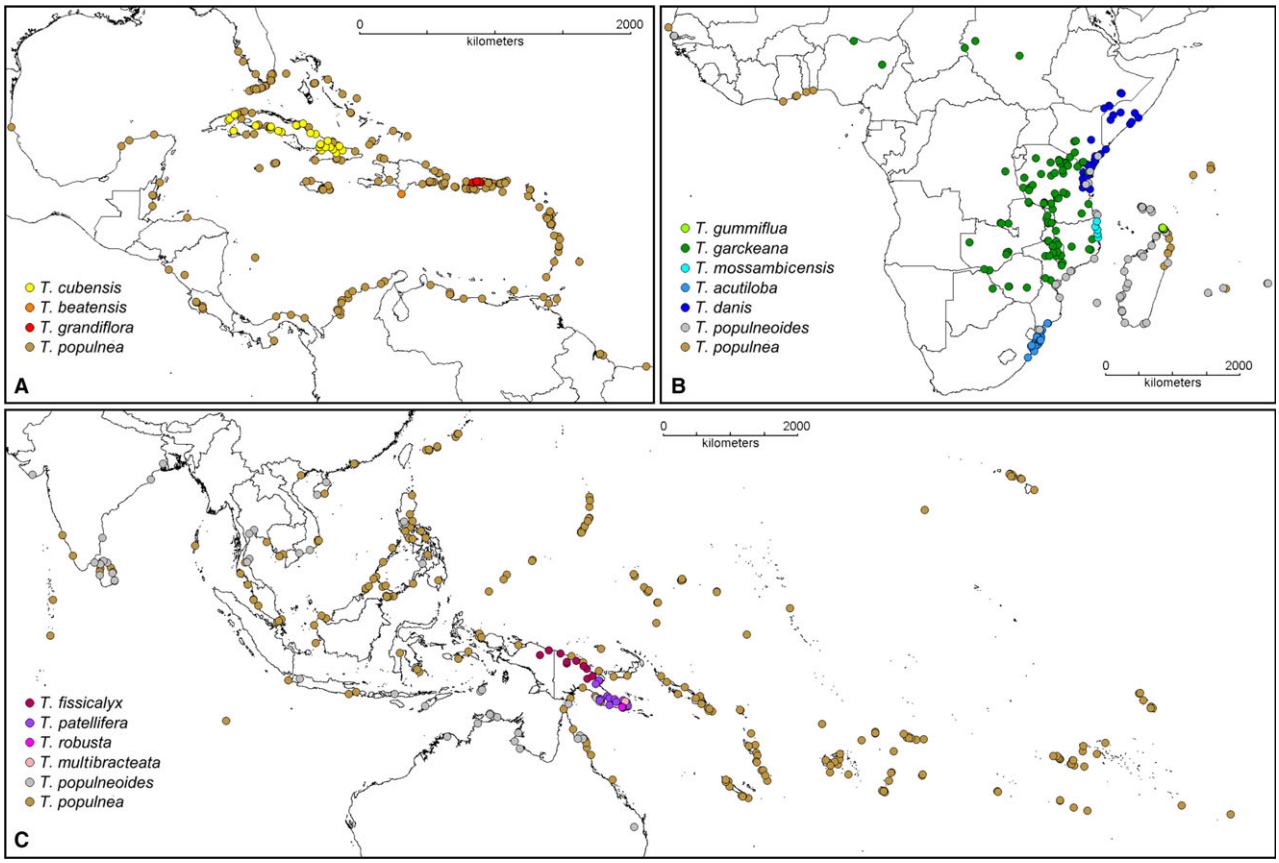

Figure 2. Geographical distribution of species of *Thespesia* section *Thespesia*: A, the New World; B, Africa; C, Southeast Asia and Oceania. Dots are localities obtained from > 1900 herbarium specimens examined.

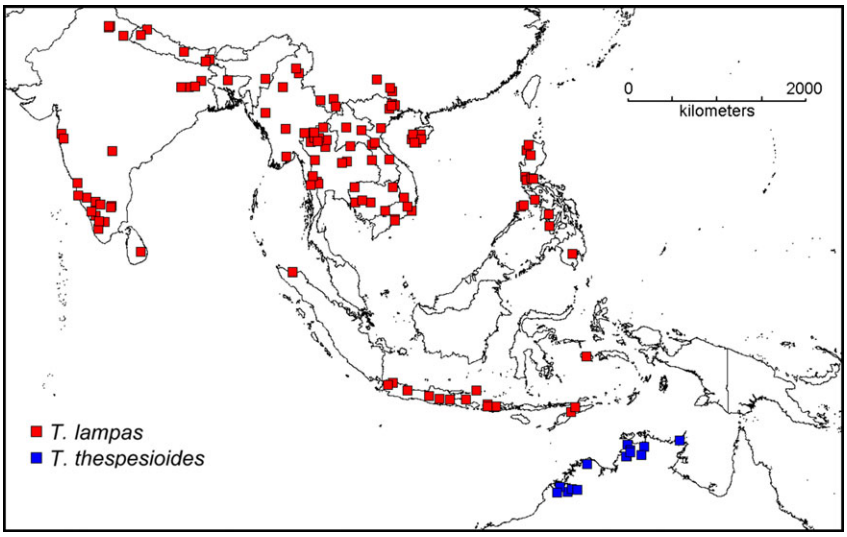

Figure 3. Geographical distribution of species of *Thespesia* section *Lampas*. Squares are localities obtained from > 300 herbarium specimens examined.

50 species from America, Africa/Arabia and Australia (Fryxell, 1992) and, due to its economic importance, has received greater attention than the other members of the tribe. In addition to *Thespesia*, the other genera are as follows. *Cienfuegosia* (Fig. 1I) comprises 30 species that are mostly Neotropical, but a few are African (Fryxell, 1997; Krapovickas, 2003). *Hampea* Schltdl. (Fig. 1J) comprises 21 species distributed from Mexico to Colombia (Fryxell, 1969, 1997). *Kokia* Lewton (Fig. 1K) comprises four species endemic to Hawaii (Fryxell, 1979; Bates, 1990). *Gossypioides* comprises two species (one from East Africa and one from Madagascar). The monotypic *Cephalohibiscus* Ulbr. is from New Guinea and the Solomon Islands and the monotypic *Lebronnecia* Fosberg (Fig. 1L) is from the Marquesas Islands of French Polynesia (Fryxell, 1979).

In addition to the above genera, one more has been suggested to belong to Gossypieae. *Thepparatia* Phuph. was described to accommodate a single species from Thailand and is thought to be related to *Thespesia* (Phuphathanaphong, 2006). The species, however, lacks gossypol glands and has certain features that are unusual in Gossypieae, including a climbing habit and the presence of serrate leaves without nectaries. The affinities of this genus to *Thespesia* and Gossypieae are therefore worth investigating with molecular data.

The monophyly of Gossypieae was supported in a phylogenetic analysis of Malvaceae using plastid restriction site data (La Duke & Doebley, 1995), which was later corroborated by a larger study that included plastid (*ndhF* gene) and nuclear (ITS) sequences (Seelanan *et al.*, 1997). According to the latter study, *Cienfuegosia* is the earliest divergent genus (Fig. 4) and, based on a constant (clock-like) substitution rate and sequence divergence estimates for *ndhF*, it diverged from the rest of the tribe in the early Miocene, *c.* 19 Mya. The lineage including *Lebronnecia*, *Thespesia* and *Hampea* (the branching order of which was not fully resolved; Fig. 4) diverged from the rest of Gossypieae *c.* 15 Mya. *Gossypium* and its sister clade consisting of *Kokia* and *Gossypioides* split *c.* 12.5 Mya (Seelanan *et al.*, 1997). Because the position of continents in the Miocene was similar as they are today, Seelanan *et al.* (1997) suggested that diversification of the tribe has been strongly influenced by long-distance, oceanic dispersal.

In the most recent phylogenetic analysis of Gossypieae, only three *Thespesia* spp. were sampled and these did not form a monophyletic group (Seelanan *et al.*, 1997). *Thespesia lampas* and *T. thespesioides* (Benth.) Fryxell, both included in *Thespesia* section *Lampas* (Ulbr.) Borss.Waalk. (Fig. 3), were more closely related to *Lebronnecia* and *Hampea* than to *T. populnea* (Fig. 4). Section *Lampas* comprises the species with lobed leaves and dehiscent fruits with many seeds per locule (van Borssum Waalkes, 1966). It initially included the only species of *Cephalohibiscus* [as *Thespesia peekelii* (Ulbr.) Borss.Waalk.]. *Thespesia* section *Thespesia*, which includes *T. populnea* and the remaining species, is characterized by entire or (less commonly) lobed leaves and indehiscent or (less commonly) dehiscent fruits with two to four seeds per locule (van Borssum Waalkes, 1966; Fryxell, 1979). Fryxell (1979) noticed the apparent discontinuity between the two sections and indicated the possibility of treating them as different genera, but at that time he considered that

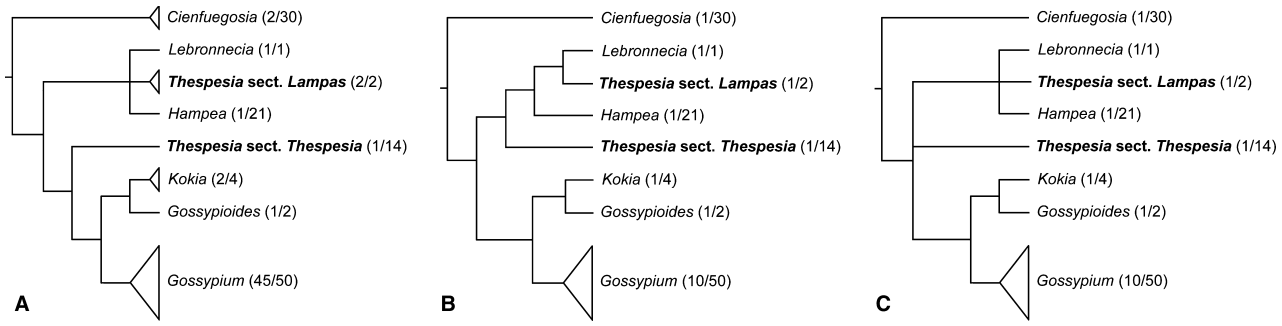

Figure 4. Schematic summary of the relationships among genera of Gossypieae based on ITS and *ndhF* sequences (modified from Seelanan *et al.*, 1997): A, phylogenetic tree based on ITS sequences; B, tree based on *ndhF* sequences; C, tree based on the combined ITS + *ndhF* dataset. The placement of the monotypic *Cephalohibiscus* is unknown. Numbers indicate the number of species included in this earlier study vs. the total number of species in each genus .

the evidence was not 'strong enough to warrant doing so' (Fryxell, 1979: 101).

The circumscription of *Thespesia* has historically varied considerably based on the weight that different authors have given to characters such as fruit dehiscence, number of seeds per locule, calyx shape, and persistence and number of involucral bracts. Taken in a broad sense, *Thespesia* comprises 16 species (*Thespesia trilobata* Baker f. is included here within the synonymy of *T. garckeana*), but almost half of these have at some time been assigned to several small genera: *Atkinsia* R.A.Howard, *Ulbrichia* Urb., *Montezuma* DC., *Thespesiopsis* Exell & Hillc., *Azanza* Alef. and *Shantzia* Lewton (Alefeld, 1861; Lewton, 1928; Howard, 1949; Exell & Hillcoat, 1954). As currently circumscribed, *Thespesia* cannot be defined by clear morphological synapomorphies, but rather a combination of several features. Its members are shrubs or more commonly trees, with three to many, linear or subulate involucral bracts, a truncate to five-lobed calyx and a three- to five-celled, oblate to elongated, dehiscent or indehiscent capsule, with few to many glabrous or pubescent seeds.

In this study we reconstruct the evolutionary relationships of *Thespesia*, within the genus and among the remaining genera of Gossypieae. We address the delimitation of the genus, investigate the phylogenetic affinities of the poorly known monotypic *Cephalohibiscus* and *Thepparatia*, and elucidate the biogeographical history of the genus emphasizing Antillean species using a relaxed molecular clock approach with fossils for calibration.

MATERIAL AND METHODS

Taxon sampling

We obtained leaf tissue from 13 of the 16 currently accepted *Thespesia* spp. using herbarium specimens or silica-gel-dried samples (Appendix 1). Tissue from herbarium specimens was received from various institutions upon request or was taken directly by the authors with permission from the corresponding herbaria. Vouchers of freshly collected samples were deposited mainly at the herbarium of the University of Puerto Rico in Río Piedras (UPRRP), but also at the herbarium of the National Botanic Garden of Cuba (HAJB), the herbarium of the National Botanic Garden Dr. Rafael M. Moscoso in Dominican Republic (JBSD), the herbarium of the National Institute of Biodiversity in Costa Rica (INB) and the herbarium of the Universidad Pedagógica y Tecnológica de Colombia (UPTC).

Whenever possible, *Thespesia* spp. were represented by material from more than one locality. The three species that could not be sampled belong to *T.* section *Thespesia* and were *T. mossambicensis* (Exell & Hillc.) Fryxell (endemic to Mozambique) and *T. multibracteata* Borss.Waalk. and *T. robusta* Borss.-Waalk. (both from Papua New Guinea). *Thespesia mossambicensis* is closely related to *Thespesia danis* Oliv. (which occurs in Tanzania, Kenya, Somalia and Ethiopia), whereas the two Papuan species are morphologically similar to the other two sampled New Guinea species with which they form a well-defined and distinctive group.

We also sampled the Thai endemic genus *Thepparatia* and the following genera of Gossypieae: *Hampea* (four species), *Cienfuegosia* (one species), *Kokia* (one species) and the monotypic *Cephalohibiscus* and *Lebronnecia*. Additional sequences of *Cienfuegosia*, *Kokia*, *Gossypium*, *Gossypioides* and *Thespesia*, mostly derived from the works of Seelanan *et al.* (1997) and Cronn *et al.* (2002), were incorporated from GenBank (Appendix 1).

Outgroups were represented by one species of *Alyogyne* Alef., the sister genus of Gossypieae (Pfeil *et al.*, 2002; Baum *et al.*, 2004), and by species

from eight genera of Malveae (*Abutilon* Mill., *Alcea* L., *Bastardia* Kunth, *Malope* L., *Malva* L., *Modiola* Moench, *Sida* L. and *Sphaeralcea* A.St.-Hil.), five from Hibisceae (*Hibiscus* L., *Malachra* L., *Malvaviscus* Fabr., *Pavonia* Cav. and *Urena* L.) and three from early-diverging genera of core Malvoideae (*Howittia* F.Muell., *Lagunaria* G.Don and *Radyera* Bullock) (Baum *et al.*, 2004). Sequences of *Alyogyne*, *Alcea*, *Malope*, *Modiola*, *Sphaeralcea*, *Malva* (except for the *matK*/*trnK* region), *Howittia*, *Lagunaria* and *Radyera* were obtained from GenBank (Appendix 1); the others were generated in this study.

Marker selection, DNA extraction, amplification and sequencing

We selected five markers, three plastid and two nuclear, for amplification and sequencing: the *matK* gene with the flanking *trnK* intron, the *ndhF* gene, the *rpl16* intron, the ribosomal ITS region and the *CesA1* gene (exons 3–5 and their associated introns). These regions have been previously used in studies involving members of Gossypieae (Seelanan *et al.*, 1997; Cronn *et al.*, 2002) and have been shown to be highly valuable for phylogenetic inference at the species level in Malvoideae (Pfeil *et al.*, 2002; Fuertes Aguilar, Fryxell & Jansen, 2003; Small, 2004; Tate *et al.*, 2005; Tate, 2011; Koopman & Baum, 2008; Escobar García *et al.*, 2009; Schneider *et al.*, 2011).

Genomic DNA was extracted from dried leaf tissue using the DNeasy plant mini kit (Qiagen). For herbarium samples the following protocol modifications were made: 500 μL buffer AP1 instead of 400 μL, an incubation time of 15–20 min instead of 10 min, 50 μL buffer AE instead of 100 μL and an elution time of 10–20 min instead of 5 min. These modifications, taken from Drábková, Kirschner & Vlček (2002), improved the DNA yield from old samples.

Due to the low quality of DNA extracted from most herbarium samples, the three plastid regions were amplified and sequenced in fragments of 600–800 bases each, which were subsequently assembled manually. The *ndhF* gene (~2100 bp) and the *matK*/*trnK* region (~2500 bp) were divided into three overlapping fragments and the *rpl16* intron (~1200 bp) was divided into two. The primer pairs used are shown in Table 1.

The ITS and plastid sequences were obtained for all species, except for one fragment of *ndhF* that could not be amplified for *T. fissicalyx* Borss.Waalk. For *CesA1*, only sequences of members of Gossypieae, the orthology of which was confirmed by comparison with previously published sequences, were included. *CesA1* sequences obtained for outgroup species (Malveae and Hibisceae) were highly discordant and were thus excluded from the dataset.

Amplification was performed with standard polymerase chain reaction (PCR). Cycling conditions for the three fragments of *trnK*/*matK* consisted of 4 min initial denaturation at 95 °C followed by 35 cycles of 1 min denaturation at 95 °C, 1 min annealing at 48 °C and 1 min extension at 72 °C. The final extension consisted of 10 min at 72 °C. Amplification of *rpl16* and the first and second fragments of *ndhF* consisted of 2 min initial denaturation at 94 °C followed by 30 cycles of 1 min denaturation at 94 °C, 30 s annealing at 54 °C and 1 min extension at 72 °C. The final extension consisted of 5 min at 72 °C. For fragment three of *ndhF* the annealing temperature was lowered to 48 °C. Cycling conditions for the ITS region consisted of 2 min of initial denaturation at 94 °C followed by 34 cycles of 20 s denaturation at 95 °C, 30 s annealing at 55 °C and 20 s extension at 72 °C, with a final extension of 2 min at 72 °C. For *CesA1*, the 2 min of initial denaturation at 94 °C were followed by 35 cycles of 30 s denaturation at 94 °C, 2 min annealing at 56 °C and 2 min extension at 72 °C. The final extension consisted of 6 min at 72 °C.

Successful reactions were cleaned with ExoSap (Fermentas) following the manufacturer's instructions. Sequencing was performed in both forward and reverse directions with the same primers used for amplification. The reactions were run on an ABI 3130xl Genetic Analyzer (Applied Biosystems) at the Sequencing and Genomics Facility (SGF) of the University of Puerto Rico, Rio Piedras. The resulting sequences were edited with Sequencher 4.8 (Gene Codes Corp.) and deposited in GenBank under accession numbers KT966901–KT967074 (Appendix 1).

Sequence alignment and phylogenetic analyses

We aligned sequences using Saté 2.2.7 with its default settings (Liu *et al.*, 2012; Liu & Warnow, 2014), and then manually adjusted in Mesquite 2.75 (Maddison & Maddison, 2011). Certain difficult fragments were realigned using the 'Realign Region' option in Geneious 8.1.3 (Biomatters, Ltd). The five sets of aligned sequences were then concatenated and analysed in combination. We submitted this combined matrix to Dryad http://dx.doi.org/10.5061/dryad.sv164.

Relationships among species were inferred with maximum-likelihood (ML) and Bayesian inference. We determined the best-fit partitioning scheme and models of molecular evolution using PartitionFinder 1.1.0 (Lanfear *et al.*, 2012). Protein-coding regions were partitioned by codon position in all analyses. The correct reading frame in these regions was determined with Geneious 8.1.3 and the NCBI Open

Table 1. Primers used in this study

Name	Sequence (5′–3′)	Reference(s)
matK/trnK		
trnK3914F	GGGGTTGCTAACTCAACGG	Cronn *et al.* (2002) modified from Johnson & Soltis (1994)
matK966R	TCGATTTATTTAAACCATGCTCA	This study
matK905F	CCCACTTATTTTTCGGGAGT	This study
matK1833R	AGCCCAGAAAGTCGAGAGAA	This study
matK1755F	TCCTATTTTTACCTGTGGTCTCA	Sánchez Andrade (2005)
trnK2R	AACTAGTCGGATGGAGTAG	Johnson & Soltis (1994)
ndhF		
F1	GAATATGCATGGATCATACC	Seelanan *et al.* (1997)
R708	CCATGGCATCAGGTAACCAT	This study
F679	TATTATTTGCCGGCGCAGT	This study
R1460	TGGCTGCTCGTTGTTATTCA	This study
F1421	TGGGGTAAAGAAGAGCAAAAA	This study
R2110	CCCCCTAYATATTTGATACCTTCTCC	Olmstead & Sweere (1994)
	TGTAATGCCTACTCCATTTGTAATTC (alternative)	This study
rpl16		
F71	GCTATGCTTAGTGTGTGACTCGTTG	Jordan, Courtney & Neigel (1996)
R631	CTGGTTTGTTCCGCCATC	This study
F560	GCATTAATCGAGAAGCGATG	This study
R1516	CCCTTCATTCTTCCTCTATGTTG	Kelchner & Clark (1997)
ITS		
ITS5	GGAAGTAAAAGTCGTAACAAGG	White *et al.* (1990)
ITS4	TCCTCCGCTTATTGATATGC	White *et al.* (1990)
CesA1		
CelA1scpF	CATTTGARCAAGTCTCAGGTATTGTT	Cronn *et al.* (2002)
CelA1R	GGGAACTGATCCAACACCCAGGA	Cronn, Small & Wendel (1999), Cronn *et al.* (2002)

Reading Frame Finder tool (http://www.ncbi.nlm.nih.gov/gorf/gorf.html).

ML analyses were performed in RAxML (Stamatakis, 2006, 2014) using the application RaxmlGUI 1.3.1 (Silvestro & Michalak, 2012). Eight partitions were specified as follows (as determined by PartitionFinder): partition 1: *trnK5* intron and the 1st and 2nd codon positions of exon 4 of *CesA1*; partition 2: the 1st and 2nd codon positions of *matK* and *ndhF* and the 2nd codon position of exon 5 of *CesA1*; partition 3: the 3rd codon position of *matK* and *ndhF*; partition 4: *trnK3* and *rpl16* introns and the 2nd codon position of exon 3 of *CesA1*; partition 5: the ribosomal 18S and 26S and the 1st codon position of exon 5 of *CesA1*; partition 6: ITS1 and ITS2; partition 7: the ribosomal 5.8S and the 3rd codon position of exon 3 of *CesA1*; and partition 8: each of the three introns of *CesA1*, the 1st codon position of exon 3 of *CesA1* and the 3rd codon positions of exons 4 and 5 of *CesA1*. We conducted ten runs using GTR as the model of nucleotide substitution to find the best ML tree. A standard ('slow') bootstrap analysis was performed with 1000 replicates to estimate clade support.

The Bayesian analysis was implemented in MrBayes 3.2.3 (Ronquist *et al.*, 2012) specifying the DNA substitution model selected with PartitionFinder for each partition: K81uf+G for partition 1, GTR+G for partition 2, TVM+G for partition 3, F81+I+G for partition 4, JC for partition 5, TrNef+I+G for partition 6, K80+1 for partition 7 and HKY+G for partition 8. We performed four runs with five million generations with four chains, sampling trees every 1000 generations. The first 20% of trees was discarded and a 50% majority rule consensus tree was reconstructed from the remaining post burn-in trees. The program Tracer 1.6.0 (Rambaut *et al.*, 2014) was used to verify that all runs reached stationarity and converged on the same distribution.

Dating analyses and fossil selection

To estimate divergence times, we conducted a Bayesian relaxed molecular clock analysis with the combined dataset in BEAST 2.3.0 (Drummond & Rambaut, 2007; Drummond *et al.*, 2012; Bouckaert *et al.*, 2014). Partitions were unlinked and assigned

their corresponding substitution model parameters. Models that cannot be set in BEAUti were specified by editing the .xml file. The uncorrelated clock with a lognormal distribution was selected as clock model and the fossilized birth death (FBD) model (Heath, Huelsenbeck & Stadler, 2014) was selected as tree prior.

Unlike the traditional 'node calibration' or 'calibration density method' in which only the oldest fossil of a clade is used to calibrate a node, the FBD approach allows one to incorporate all reliable fossils known for a lineage, which are integrated with modern taxa to the diversification process (Heath *et al.*, 2014). Moreover, the FBD model does not require one to apply arbitrary prior distributions to the calibration nodes and seems to provide more accurate approximations of divergence times and of statistical uncertainty compared with traditional approaches (Heath *et al.*, 2014; Grimm *et al.*, 2015).

Based on a comprehensive literature review, we selected 18 fossil records (see Appendix 2 for location, age ranges, affinities and references) that can be confidently assigned to core Malvoideae (*sensu* Baum *et al.*, 2004). The oldest corresponds to the petrified wood *Hibiscoxylon nyloticum* Kräusel collected from Late Cretaceous deposits (88–65 Myr old) of Egypt (Kräusel, 1939) and Ethiopia (Beauchamp & Lemoigne, 1973; Gros, 1992). The wood of *H. nyloticum* has been related to that of the arborescent *Lagunaria patersonia* (Andrews) G.Don (Kräusel, 1939; as *Hibiscus patersonius* Andrews), one of the early-branching genera of core Malvoideae (Baum *et al.*, 2004) and thus this fossil was used to date the most recent common ancestor of the clade including *Lagunaria* and the eumalvoid clade. Next oldest fossil records include leaves of *Malvaciphyllum macondicus* M. Carvalho from the mid- to late Palaeocene (60–58 Myr old) of northern Colombia (Carvalho *et al.*, 2011) and pollen grains (species of *Malvacipollis* Harris) from the mid- and late Palaeocene of North America (Fredericksen, 1980; Gaponoff, 1984).

Many fossil pollen deposits referrable to Malvoideae have been recorded from the Eocene and later. Based on certain palynological features some of these grains can be assigned to specific tribes in the eumalvoid clade. For instance, *Malvacearumpollis mannanensis* Wood, *Malvacearumpollis bakonyensis* Nagy, *Baumannipollis chubutensis* Barreda, *Malvacipollis argentina* Zamaloa & Romero and *Malvapantocolporites silvinites* Mautino, Cuadrado & Anzótegui (Appendix 2) were all assigned to Malveae. These species were described from triaperturate or oligoaperturate grains with spines sitting on prominent basal cushions, features that are common to modern members of Malveae (Christensen, 1986). *Echiperiporites estelae* Germeraad, Hopping & Muller and several other polyporate '*Hibiscus*-type' grains (Appendix 2) were assigned to Hibisceae, a tribe characterized by having a higher number of apertures and spines lacking (or with indistinct) basal cushions (Christensen, 1986).

Malvocarpon clarum Hollick, a fossil known from the Oligocene (*c.* 29 Myr old) of Puerto Rico (Hollick, 1928; Graham, 2003), which was described from a schizocarpic fruit similar to that of *Abutilon*, was also placed in Malveae, in the clade formed by members of the *Abutilon* alliance. Fossil pollen that has been assigned to *Abutilon* has been also recovered from the same Oligocene deposits in Puerto Rico (Graham & Jarzen, 1969; Graham, 2003), which further supports the presence of members of the *Abutilon* alliance during this geological epoch.

The only valid fossil record of Gossypieae of which we are aware corresponds to recent (135 000–120 000 years old) fossil leaves of Hawaiian cotton, *Gossypium tomentosum* Nutt. ex Seem. (Woodcock & Manchester, 1998). This fossil was placed in the clade formed by the allopolyploid cottons (subgenus *Karpas* Raf.). A fossil leaf of *Hampea* has been reported from the Tertiary of Alaska (Hollick, 1936), but due to its resemblance to several other genera from different plant families and the fact that *Hampea* is a Central American genus, the identity of this record is highly questionable and thus it was not included in our analyses.

We assigned an exponential distribution to the two hyperparameters of the relaxed clock model. A mean of 10.0 was set for the *ucldMean.c* and a mean of 0.333 was assigned to the *ucldStdev.c*. The priors of the FBD model were set as follows: an exponential distribution with a mean of 1.0 was assumed for the diversification rate, a lognormal prior distribution with the offset corresponding to the age of the oldest fossil (88 Myr) was assigned to the origin time (Mean = 1.5, SD = 1.0) and a uniform (0,1) prior distribution was assumed for both turnover and sampling proportion.

Two independent runs were conducted in BEAST, each with 30 million generations and sampling trees every 1000 generations. We used the software Tracer to inspect parameter values and confirm that both runs converged. The sampled trees from both runs were combined into a single file with the program LogCombiner 2.3.0 (Rambaut & Drummond, 2014a). The first 20% of trees was discarded and a majority rule consensus tree was constructed from the remaining post burn-in trees with TreeAnnotator 2.3.0 (Rambaut & Drummond, 2014b). Because fossils were only used as calibration points for the clades of interest and no morphological information was added, we pruned them from the trees with the program

Full2ExtantConvertor.jar by A. Gavryushkina (included in the output files at http://treethinkers.org/tutorials/divergence-time-estimation-using-beast/).

Ancestral range estimation

Ancestral ranges of *Thespesia* and closely related genera were inferred with the package 'BioGeoBEARS' (Matzke, 2013a) implemented in R 3.1.3 (R Development Core Team, 2015). We used the BayArea-like and the BayArea-like + J models (Matzke, 2013b), and determined which of the two best fits our data by calculating a value for the likelihood ratio test (LRT). The BayArea-like model is a simplified likelihood interpretation of the Bayesian BayArea model of Landis *et al.* (2013). It assumes dispersal, extinction, narrow and widespread sympatry, but not vicariance (Matzke, 2013b). BayArea-like and BayArea-like + J models include the free parameters: d (dispersal or range expansion) and e (extinction or range contraction), but the BayArea-like + J model also includes the parameter J that accounts for founder-event speciation (Matzke, 2013b, 2014a).

Thespesia spp. and the genera included in the sister clade were assigned to three major geographical areas: Africa including Madagascar (area A); Southeast Asia and Oceania (area B); and America (area C). From the maximum clade credibility tree inferred with BEAST, we extracted the clade containing those genera and used it as the input tree in BioGeoBEARS. The probabilities of the ancestral states obtained with the best-fitting model were drawn as pie charts at the nodes of this tree.

Biogeographical events were counted with a biogeographical stochastic mapping analysis (Matzke, 2014b). We performed 500 stochastic mapping simulations using the best-fitting model to obtain the number of anagenetic (dispersal) and cladogenetic (sympatry and founder) events.

RESULTS

Phylogenetic relationships

In total, 174 new sequences were generated in this study (39 *matK*/*trnK*, 37 *ndhF*, 38 *rpl16*, 37 ITS and 23 *CesA1*) and 99 were incorporated from GenBank. The combined matrix included 53 species (61 accessions) and consisted of 7630 aligned characters, of which 2681 corresponded to *matK*/*trnK*, 2094 to *ndhF*, 1347 to *rpl16*, 800 to ITS and 708 to *CesA1*. No characters were excluded from analyses. The number of variable and constant sites, number of species and accessions included, and percentage of missing data for each of the five regions included in this study and the combined dataset are presented in Table 2.

ML and Bayesian analyses of the combined dataset produced trees with almost identical topologies and high support values for most of the clades. The only difference between the two analyses was the position of two outgroup species, *Alcea rosea* L. and *Sphaeralcea angustifolia* (Cav.) G.Don (Malveae), which were recovered as sisters in the ML tree. The Bayesian 50% majority rule consensus tree with posterior probabilities and bootstrap support values resulting from the ML analysis is shown in Figure 5.

As revealed in previous studies, *Thespesia* is not monophyletic. Our analyses separated the species into two groups, which correspond to the two sections of the genus, sections *Thespesia* and *Lampas*. The species of *T.* section *Thespesia* formed a strongly supported clade [bootstrap support (BS) = 100%, posterior probability (PP) = 1], which was subdivided into two clades: one comprising African species [*T. gummiflua* Capuron, *T. danis*, *T. acutiloba* (Baker f.) Exell & Mendonça and *T. garckeana*]; and the other comprising the New World and the remaining Old World species. The three Greater Antilles endemics (*T. cubensis*, *T. beatensis* and *T. grandiflora*) formed a clade that is sister to a clade formed

Table 2. Description of individual DNA regions and the combined dataset including aligned length, number of variable and constant sites, number of species included and percentage of missing data

Region	*matK*/*trnK*	*ndhF*	*rpl16*	ITS	*CesA1*	Combined dataset
Position in the aligned matrix	1–2681	2682–4775	4776–6122	6123–6922	6923–7630	1–7630
Number of aligned sites	2681	2094	1347	800	708	7630
Number of constant sites	2111	1759	1052	380	567	5869
Number of variable sites	570	335	295	420	141	1761
% of variable sites	21.3	16	21.9	52.5	19.9	23.1
Number of species (and accessions) included	52 (60)	53 (61)	49 (57)	50 (58)	29 (37)	53 (61)
% of missing data	2.2	1.1	6.3	7.7	39.1	6.6

by the New Guinea species plus the widely distributed *T. populnea* and *T. populneoides* (Roxb.) Kostel. (Fig. 5).

The two species of *T.* section *Lampas* are also placed together in a well-supported clade (BS = 100%, PP = 1) that includes *Lebronnecia*, *Cephalohibiscus* and *Hampea*. *Lebronnecia* is the sister genus of *T.* section *Lampas*, whereas the other monotypic genus, *Cephalohibiscus*, is sister to *Hampea* (Fig. 5). The encompassing clade consisting of these four taxa plus *Thespesia* section *Thespesia* is also strongly supported (BS = 100%, PP = 1).

The position of the remaining genera of Gossypieae is congruent with previous phylogenetic studies: *Cienfuegosia* is the earliest-branching genus and the monophyletic *Gossypium* has its closest relationship with *Gossypioides* and *Kokia*. The *Gossypium–Gossypioides–Kokia* clade is in turn sister to the one comprising the above-mentioned genera (Fig. 5).

The Thai endemic *Thepparatia* was not related to any member of Gossypieae. Instead, this monotypic genus was recovered in the clade formed by species of Hibisceae (Fig. 5).

Divergence times

The maximum clade credibility tree resulting from the partitioned analysis of the five-marker dataset using the FBD approach with 18 fossil records is shown in Figure 6. The analysis places the separation of *T.* section *Thespesia* from its sister lineage (formed by *Lebronnecia*, *T.* section *Lampas*, *Cephalohibiscus* and *Hampea*) in the Oligocene, *c.* 30 Mya [23–38 Mya, 95% highest posterior density interval (HPD)], but extant species did not start to diverge until the late Miocene. The African lineage separated first from the ancestor of all other species *c.* 11 (8–14, 95% HPD) Mya, whereas the Greater Antilles lineage diverged from the mostly Asian group about 9 (7–12, 95% HPD) Mya. With the exception of the two sampled New Guinea species, which split *c.* 3.8 Mya, the lineages leading to individual species of *T.* section *Thespesia* separated from each other between 9 and 6 Mya (Fig. 6).

The divergence of the four lineages in the sister clade of *T.* section *Thespesia* occurred in the early Miocene (Fig. 6). *Thespesia* section *Lampas* diverged from *Lebronnecia c.* 16 (10–22, 95% HPD) Mya, although the ancestor of the two extant species arose much more recently, *c.* 3.5 Mya. *Cephalohibiscus* and *Hampea* split *c.* 18 (12–23, 95% HPD) Mya. This last genus began to diversify at least 7.6 (5–11, 95% HPD) Mya.

Gossypieae originated in the early Palaeocene, *c.* 65 (56–75, 95% HPD) Mya, but extant genera began to differentiate in the mid- to late Eocene, *c.* 41 (33–50, 95% HPD) Mya. Shortly after the separation of *Cienfuegosia*, the other two main lineages in the tribe split *c.* 38 (30–46, 95% HPD) Mya. *Gossypium* separated from *Gossypioides* and *Kokia* in the Oligocene, *c.* 28 (21–35, 95% HPD) Mya. These divergence dates for Gossypieae and the included lineages are much older than previously estimated.

Our analysis also indicated a much earlier origin for the eumalvoids and the other two tribes in the clade. This group originated in the late Cretaceous, about 82 (73–89, 95% HPD) Mya, and began to diversify approximately 75 (66–83, 95% HPD) Mya. The crown groups of Malveae and Hibisceae are at least 55 (45–66, 95% HPD) and 41 (31–52, 95% HPD) Myr old, respectively (Fig. 6).

Ancestral range estimation

Of the two biogeographical models tested, the one that best fitted the phylogenetic and geographical data for the clade comprising *Thespesia* and closely related genera was the BayArea-like + J model (LnL = −21.05 vs. −29.11, LRT = 16.13, $P = 0.00006$). The ancestral range estimation using this model indicates that *T.* section *Thespesia* most probably originated in Southeast Asia–Oceania (Fig. 7). Two founder events, one into Africa (about 11 Mya) and another into America (2 My later), gave rise to the African and the American (Antillean) species, respectively.

As expected given their present distribution, *Lebronnecia*, *T.* section *Lampas* and *Cephalohibiscus* also originated in Southeast Asia–Oceania. Another founder event into America, which preceded the other two, led to the origin of *Hampea* (Fig. 7). Asia is also the most likely origin for the encompassing clade, which includes these taxa plus *T.* section *Thespesia*.

In addition to the three founder events, the biogeographical stochastic mapping indicated two anagenetic dispersals (Table 3). These correspond to the dispersal of *T. populneoides*, most probably from Southeast Asia to Africa and of *T. populnea* from Asia to Africa and to America (Fig. 7).

DISCUSSION

Phylogenetic relationships and delimitation of *Thespesia*

The ML and Bayesian analyses of the combined five-marker dataset largely corroborated the phylogenetic relationships revealed in previous studies for the genera of Gossypieae (Seelanan *et al.*, 1997) and provided resolution for some previously unresolved relationships. Specifically, our analyses confirmed the non-monophyly of *Thespesia*, clarified the branching order of the clade including *Lebronnecia*, *Hampea* and *T.* section *Lampas*, and confirmed the sister relation-

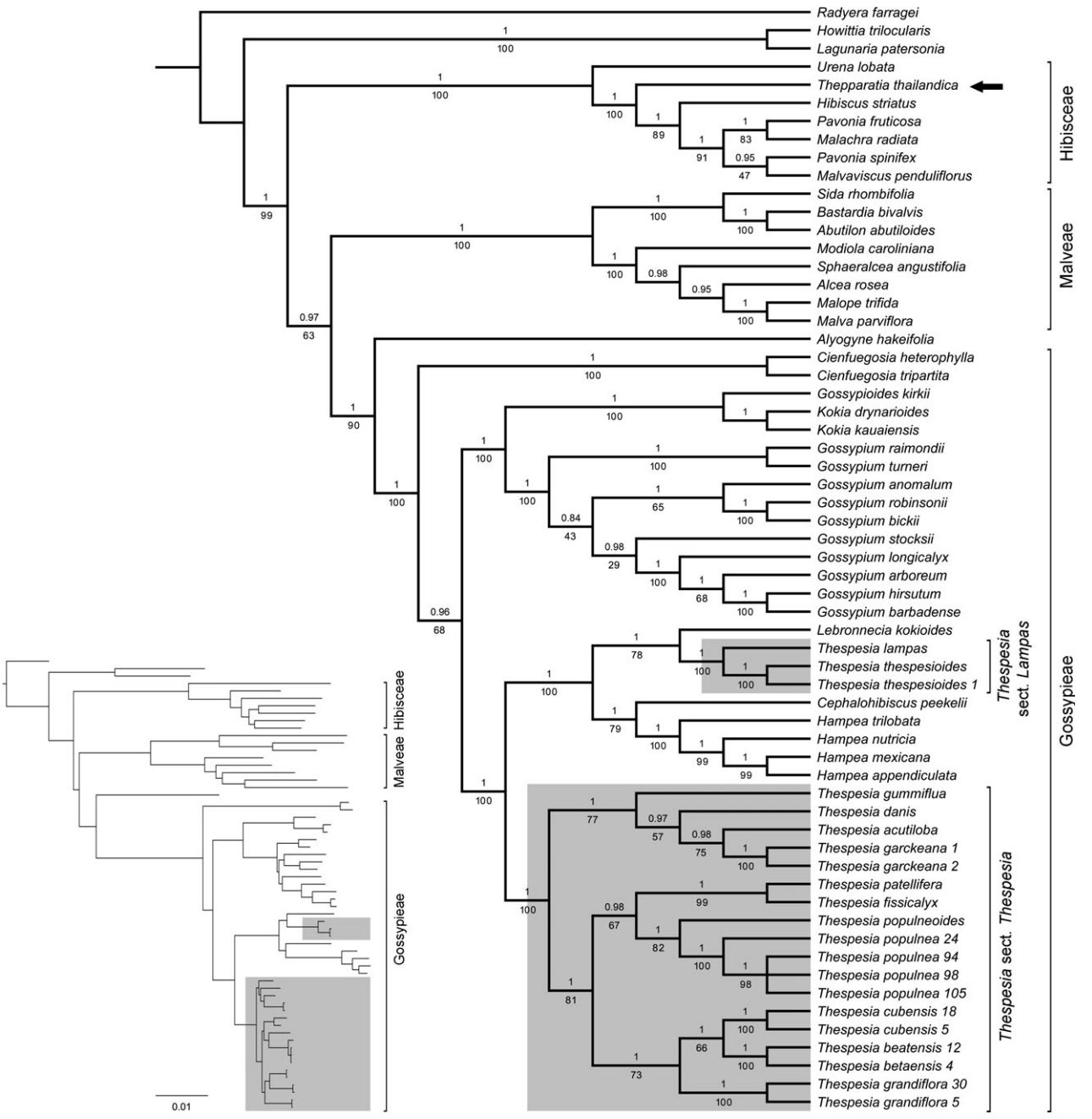

Figure 5. Bayesian 50% majority rule consensus tree of *Thespesia* and tribe Gossypieae resulting from the analysis of the concatenated dataset (*trnK/matK*+*ndhF*+*rpl16*+ITS+*CesA1*). Posterior probabilities (PP) and bootstrap support (BS) values are shown above and below branches, respectively. The topology is nearly identical to that resulting from the ML analysis, except for the position of *Alcea* and *Sphaeralcea* in tribe Malveae (BS values were omitted for the branches involved). Clades containing *Thespesia* spp. are highlighted in grey. The arrow indicates the position of *Thepparatia thailandica* in tribe Hibisceae.

ship of this group (which also includes *Cephalohibiscus*) to *T.* section *Thespesia* (Fig. 5).

Thespesia, as currently circumscribed, is biphyletic and therefore its boundaries need to be redefined. *Thespesia* should be restricted to include the 14 species of *T.* section *Thespesia*, and the two species of *T.* section *Lampas* (*T. lampas* and *T. thespesioides*) need to be transferred to a different genus. In this regard, we propose to reinstate *Azanza*, which is currently included in the synonymy of

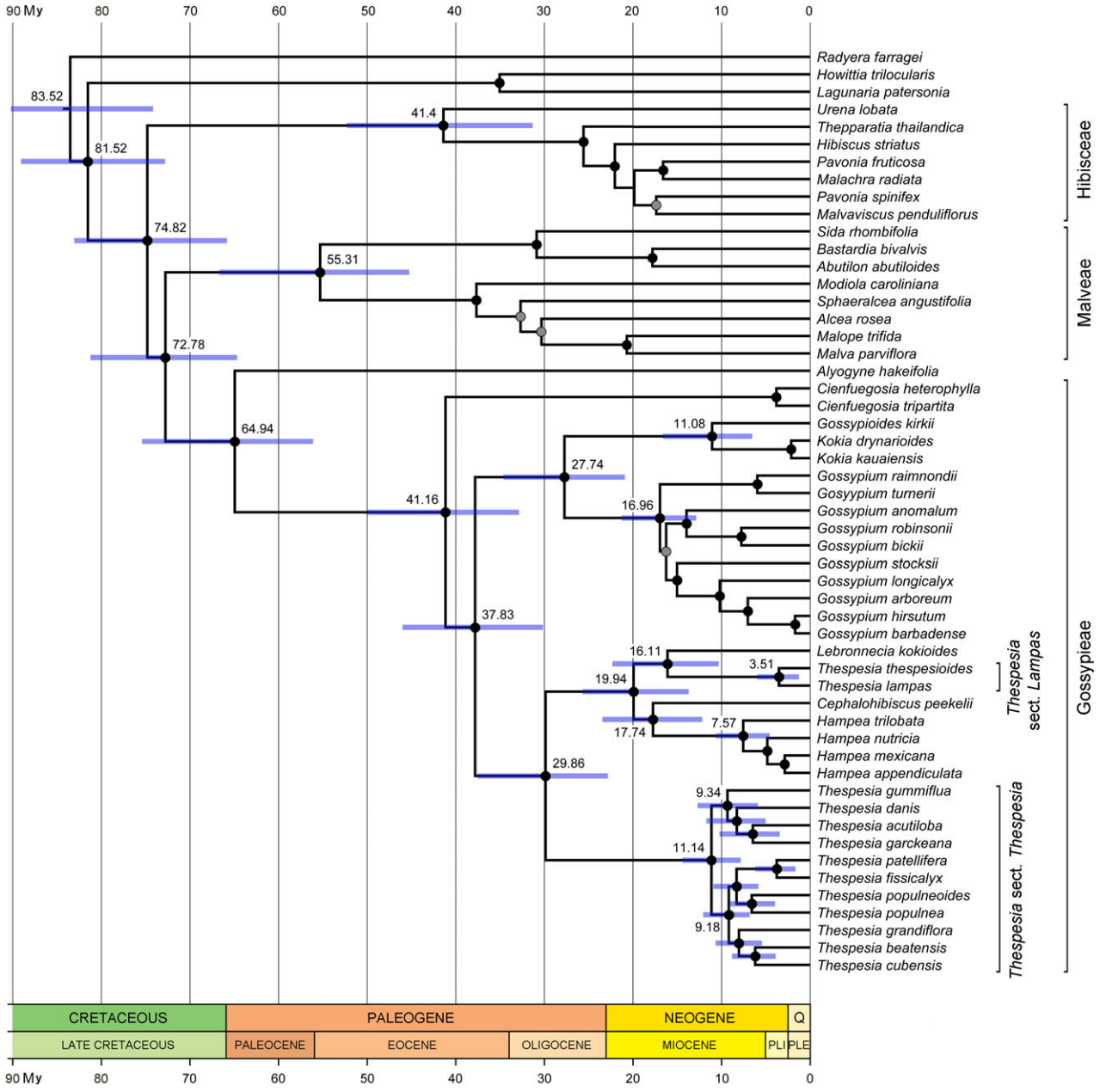

Figure 6. Chronogram of *Thespesia* and tribe *Gossypieae* resulting from the analysis of the five-marker dataset (*trnK/matK*+*ndhF*+*rpl16*+ITS+*CesA1*) using the FBD model in BEAST 2. Nodes with black circles have posterior probabilities (PP) between 0.95 and 1, whereas nodes with grey circles have PP between 0.7 and 0.949. Numbers are node ages (shown only for relevant nodes), and blue bars correspond to the 95% highest posterior density intervals (HPD) for node ages.

Thespesia, to accommodate the two species in question. *Azanza* was described by Alefeld (1861) to segregate *Hibiscus lampas* Cav. on the basis of its unbranched style (vs. branched in *Hibiscus*) and the different structure of the epicalyx and calyx. Alefeld correctly noticed that this species was in fact allied to genera of Gossypieae (*Thespesia* and *Gossypium*), but not similar enough to be included in any of them. He distinguished *Azanza* from *Thespesia* mainly by its dehiscent fruits and the numerous, glabrous seeds per locule. However, subsequent authors who favoured the inclusive circumscription of *Thespesia* largely ignored his observations, which are here substantiated with molecular data.

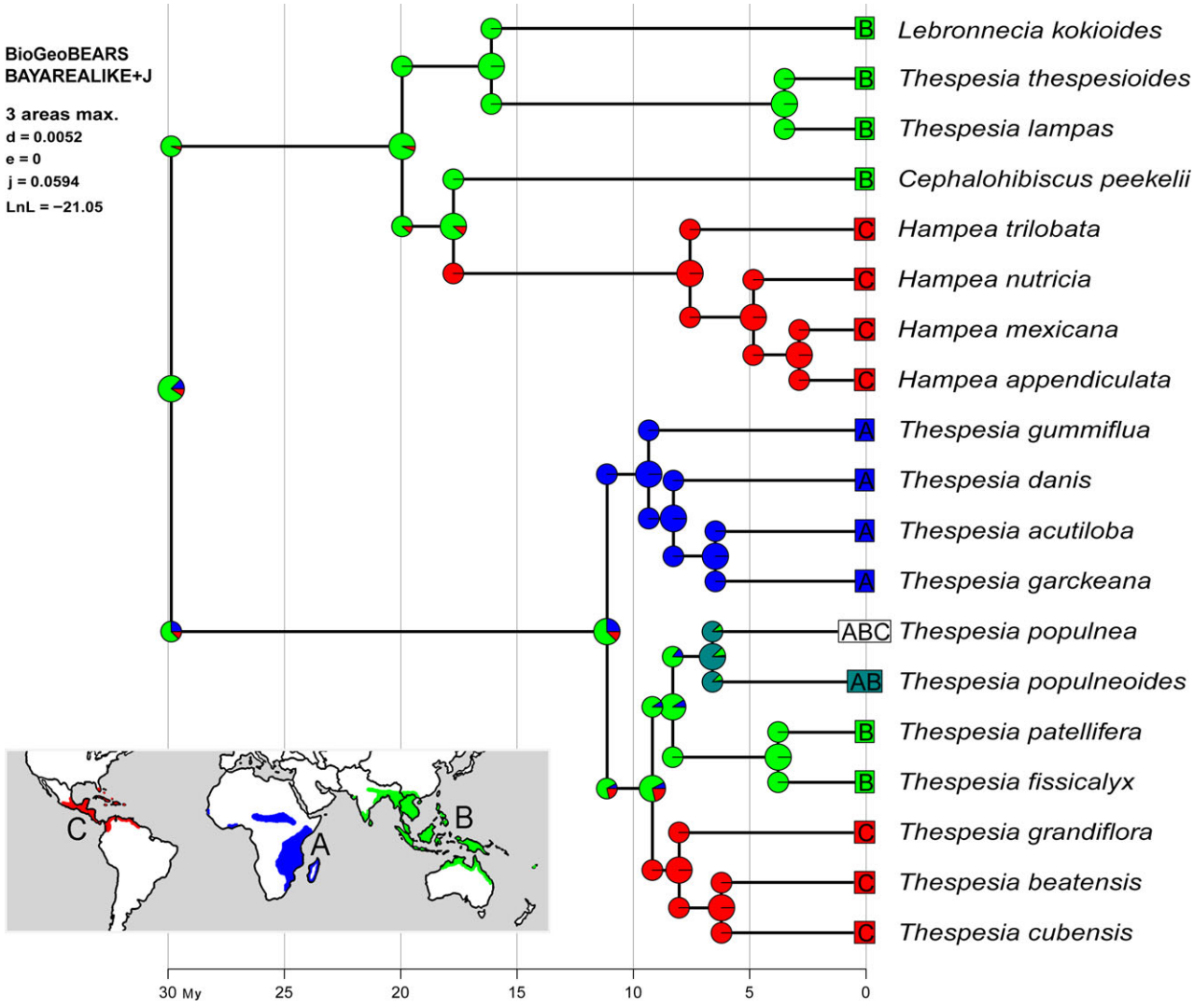

Figure 7. Ancestral ranges of *Thespesia* and related genera inferred with the package BioGeoBEARS using the BayArea-like + J model. The relative probability of each of the three ancestral states is shown as a pie chart at the node of each clade. Smaller pies at the corners represent the ancestral states immediately after the cladogenetic event. Squares at the tips show the current distribution of species. A (blue), Africa including Madagascar; B (light green), Southeast Asia and Oceania; C (red), America.

In addition to the characteristics noted above, *Azanza* can be differentiated from *Thespesia* and the remaining genera of Gossypieae by its shrubby habit, the presence of six to nine filiform, inconspicuous epicalyx bracts, a five-dentate to five-lobed calyx and the four- or five-celled ovoid-angled capsule with 8–12 relatively small seeds per locule. The sister placement of *Azanza* to *Lebronnecia* revealed in our phylogenetic analysis is unexpected. *Lebronnecia* has been traditionally related to *Hampea*, with which it shares the three-celled capsule, three epicalyx bracts and pale-coloured flowers (Fryxell, 1979). With the possible exception of the five-toothed calyx, we are not aware of morphological synapomorphies for the *Azanza–Lebronnecia* clade.

The segregation of *Azanza* makes *Thespesia* a much more cohesive, readily defined genus characterized by an arborescent habit, linear epicalyx bracts, a truncate (or at most five-denticulate) calyx and a (four-) five-celled, oblate to subglobose, indehiscent (or rarely dehiscent) capsule with one to four seeds per locule. The indehiscent capsule, which can be hard and woody to somewhat fleshy or coriaceous, is an exclusive trait of *Thespesia* and constitutes a synapomorphy for the genus (Fig. 8). It is present in all species except in the Antillean *T. beatensis* and the African *T. garckeana* and *T. gummiflua* (Fig. 8). In *T. populneoides*, only the exocarp splits into five valves, whereas the internal walls remain indehis-

Table 3. Biogeographical event counts for *Thespesia* and related genera from 500 biogeographical stochastic mappings under the BayArea-like + J model (ML parameters: $d = 0.0052$, $e = 0$, $j = 0.0594$)

Biogeographical event	Mean of event counts	SD of event counts
Anagenetic dispersal	2.13	0.33
Narrow sympatry	0	0
Subset sympatry	0	0
Vicariance	0	0
Founder event	3.54	0.63
TOTAL events	5.66	0.74

cent. This intermediate condition suggests that the dehiscence in *Thespesia* is secondarily acquired.

In contrast to the dehiscent capsule, in which the dispersal element is the seed, the indehiscent capsules in *Thespesia* are variously modified to be dispersed as a unit by different agents. In *T. populnea*, for example, the indehiscent capsules are buoyant and dispersed by sea currents. In *T. grandiflora*, the hanging, fleshy capsules are dispersed by bats (Francis, 1999), and the red fruits of *T. acutiloba* and *T. mossambicensis* are most probably attractive to birds.

Our study also clarified the placement of the monotypic *Cephalohibiscus* and *Thepparatia*. *Cephalohibiscus* was the only genus of Gossypieae for which phylogenetic information was lacking. We found that it is more closely related to *Hampea* than to any other genus of the tribe. *Cephalohibiscus* was traditionally associated with *Thespesia* on the basis of its five-celled fruits with many hairy seeds (Fryxell, 1968, 1979) and it was even included in this genus by van Borssum Waalkes (1966). Morphological synapomorphies for the *Cephalohibiscus–Hampea* clade are difficult to delineate, but perhaps the truncate calyx, white flowers lacking a dark throat and the patent to reflexed petals are appropriate.

The placement of *Thepparatia thailandica* Phuph. in Hibisceae was not completely unexpected given its morphology. Our sampling of Hibisceae, however, was insufficient to determine the closest affinities of the genus in this diverse tribe. *Thepparatia* was recovered as sister to a clade comprising *Hibiscus*, *Pavonia*, *Malachra* and *Malvaviscus*. It shares with

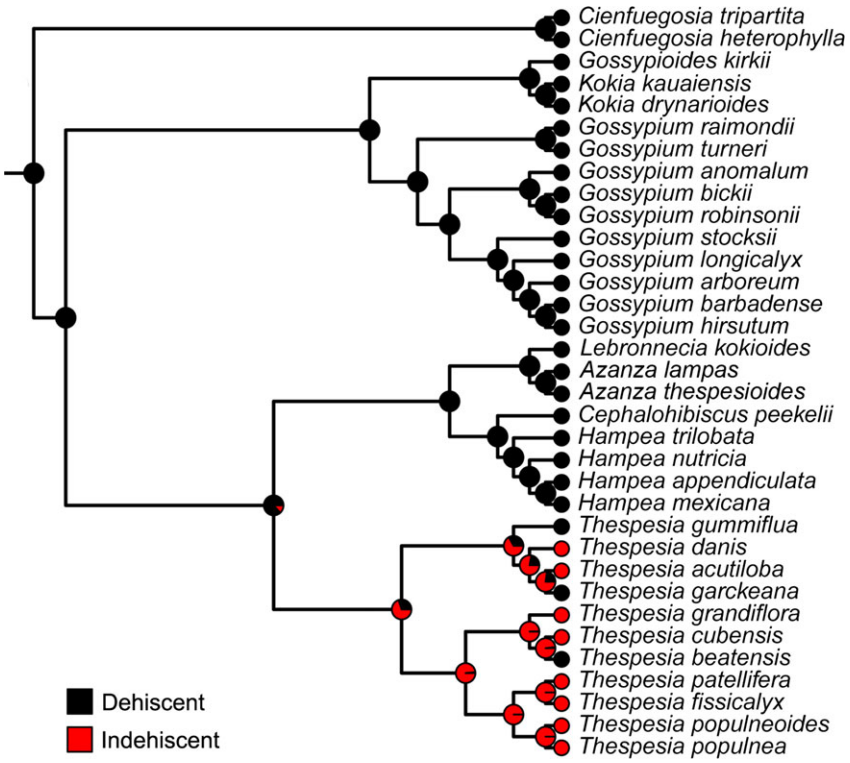

Figure 8. Ancestral state reconstruction for fruit dehiscence in tribe Gossypieae using the re-rooting method of Yang, Kumar & Nei (1995) under an 'ER' model implemented in R with the phytools function *rerootingMethod* (Revell, 2012). The state character for *T. populneoides* was assumed to be indehiscent because the endocarp remains intact and only the exocarp splits. The indehiscent fruit is an exclusive trait of *Thespesia s.s* and constitutes a synapomorphy for the newly defined genus.

these genera the deeply five-lobed calyx, a persistent epicalyx of more than five bracteoles and the pentacarpelar ovary. The style in *Thepparatia*, however, has been described as undivided (Phuphathanaphong, 2006), a condition that is not present in Hibisceae. The style was probably not fully developed in the flowers used for description, as it was also included in the staminal tube. Protandrous flowers are not unusual in Malvaceae.

Divergence times and historical biogeography

Our dating analyses and ancestral range estimates indicate that *Thespesia s.s.* originated *c.* 30 Mya in Southeast Asia and began to diversify *c.* 11 Mya with the separation of the African lineage. The current distribution of the five African endemics on the east side of the continent (Fig. 2B) is indeed consistent with an ancestral migration from Southeast Asia westwards across the Indian Ocean.

The American (Antillean) lineage diverged *c.* 9 Mya and appears to have migrated across the Pacific to the Greater Antilles possibly via the Pacific North Equatorial Countercurrent. This implies that the ancestor must have passed through the Central American Seaway, which connected the Pacific and the western Atlantic oceans until its closure, commonly regarded as occurring in the Plio-Pleistocene, *c.* 3.0–3.5 Mya (Keigwin, 1978, 1982; Duque-Caro, 1990; Burton, Ling & O'Nions, 1997; Leigh, O'Dea & Vermeij, 2014). However, recent geochronological studies suggest that the seaway closed much earlier, about 13–15 Mya (Montes *et al.*, 2015). Although this timing precedes the estimated split between the Asian and the Antillean lineage of *Thespesia* by 4–6 Myr, water exchange between the two sides of the Panamanian Isthmus probably continued for several million years along shallow, intermittent channels (Bacon *et al.*, 2015; Montes *et al.*, 2015). These pathways would have allowed the movement of floating propagules and shallow-water-dwelling organisms between both water bodies until the full closure of the isthmus sometime within the last 10–3 Myr (Bacon *et al.*, 2015). In *Thespesia*, the propagules were most probably indehiscent, buoyant fruits (Fig. 8) capable of withstanding prolonged exposure in saltwater.

By the time the Antillean ancestor diverged from the Asian lineage *c.* 9 Mya, the northern Greater Antilles were already separated (Graham, 2003; Iturralde-Vinent, 2006). Additionally, it is probable that portions of these were still submerged, so rather than three main islands there were possibly a greater number of smaller emergent land areas (Iturralde-Vinent, 2006). It is difficult to ascertain to which of these islands the common ancestor of the Greater Antillean endemics may have arrived. However, based on fruit morphology, the present distribution of the three species and palaeogeographical information, we infer that the ancestor most probably reached the south-eastern part of Cuba. This assumption is based on the following. (1) The ancestor should have had indehiscent oblate, woody fruits similar to those present in the Asian clade. *Thespesia cubensis* is the only Antillean endemic species with such fruits, whereas *T. grandiflora* has fleshy, animal-dispersed fruits and *T. beatensis* has dehiscent fruits. (2) The south-eastern part of Cuba remained subaerial throughout the Miocene and thereafter and had a much larger land area compared to other islands such as Puerto Rico. In addition, most of the southern Peninsula of Barahona in Hispaniola, including Beata Island (where *T. beatensis* is endemic), was submerged during this period of time (Iturralde-Vinent & MacPhee, 1999; Iturralde-Vinent, 2006).

The lineage leading to *T. grandiflora* diverged from the ancestor of *T. cubensis* and *T. beatensis* about 8 Mya (Fig. 6), and would have migrated from eastern Cuba to Puerto Rico. *Thespesia grandiflora* is morphologically and ecologically the most divergent of the three Antillean endemics (Fig. 1C), which is consistent with a longer time of isolation. A second over-water migration should have occurred from Cuba to southern Hispaniola to give rise to *T. beatensis*. Although prevailing surface currents in the Caribbean would make the eastward movement unlikely, drifting buoy data show considerable variation in space and time in the circulation patterns in the region (Gyory, Mariano & Ryan, 2016), some of it in the form of eddies which could float propagules from Cuba to the other Antilles.

Thespesia beatensis diverged from *T. cubensis* *c.* 6 Mya (Fig. 6). By this time, Beata Island was still underwater and remained so probably until the Quaternary, *c.* 2 Mya (Iturralde-Vinent & MacPhee, 1999). This implies that *T. beatensis* (or an extinct ancestor) should have had a wider distribution in southern Hispaniola or elsewhere. At some point it became extinct on the larger islands, surviving only on Beata Island.

The two most widely distributed *Thespesia* spp. are the hydrochorous *T. populnea* and *T. populneoides* (Figs 2, 7). Our analyses reveal that they most probably originated in Southeast Asia–Oceania from where they spread via water dispersal to Africa and America. *Thespesia populneoides* occurs along the coasts of the Indian Ocean and tropical western Pacific. In Africa, it is a relatively common seashore species along the east coast including Madagascar and nearby islands, but, except for a few collections from Senegal, it is absent from the Atlantic coast

(Fig. 2B). This distribution pattern agrees with a westward expansion of the species from Southeast Asia across the Indian Ocean. The collections from Senegal are all from the Delta du Saloum National Park [*Lykke et al. 616* (AAU, NY), *Madsen 2794* (AAU), *Madsen et al. 3034* (AAU)], an area dominated by savannas and mangroves, but also impacted by human activity (Lykke, 1994). We do not have an explanation for the presence of this species in this part of Africa other than human intervention. *Thespesia populneoides*, like *T. populnea*, is valued for wood and medicinal uses and may have been introduced. Indeed, several unambiguously exotic tree species of ethnobotanical value (e.g. *Anacardium occidentale* L., *Azadirachta indica* A.Juss and *Tamarindus indica* L.) also occur in the Park, particularly where villages have been abandoned (Lykke, 1994). Another collection from West Africa [Benin: *Chevalier 23087* (P)] is probably also from a cultivated plant.

The nearly pantropical *T. populnea* is most common along the shores of the tropical Pacific Islands where it may have been spread by their inhabitants (Friday & Okano, 2006). It is also common in the West Indies and along the continental coasts of the Caribbean. By contrast, it is sparsely distributed along the coast of mainland Africa (Fig. 2B). Unlike *T. populneoides*, this species is absent from the east coast of Africa, although it can be found in Madagascar, Mauritius, the Seychelles and other islands of the Indian Ocean. This distribution suggests that *T. populnea* has also expanded its range westwards across the Indian Ocean, but to a lesser extent than *T. populneoides*. The greatest range expansion of *T. populnea* occurred in the opposite direction, across the Pacific to Central America and the Caribbean. The relatively isolated and presumably native populations along the west coast of Africa are most probably due to recent dispersal from America across the Atlantic. Population genetic studies interpreted in the light of ocean currents are necessary to explore this hypothesis, although human intervention cannot be ruled out.

Overall, our estimates of divergence dates for genera of Gossypieae are nearly twice as old as those reported by Seelanan *et al.* (1997), who used *ndhF* sequence divergence estimates assuming a strict molecular clock. To calibrate the clock, Seelanan *et al.* (1997) used a mean rate of 5×10^{-10} substitutions per site per year, which was calculated from general rate estimates for plastid genes compiled by Palmer (1991). However, our analysis of the combined dataset in BEAST indicated no support for a strict molecular clock (*ucldStdev* = 0.36, 0.28–0.46 95% HPD). A separate analysis for the *ndhF* sequences also revealed moderate levels of variation across branches (*ucldStdev* = 0.37, 0.20–0.55 95% HPD), indicating that the evolution of this gene is not strictly clock-like. The assumption of a strict molecular clock and the use of a rough estimate of substitution rate as a method of calibration may have contributed to the underestimated divergence times reported by Seelanan *et al.* (1997).

Even though our estimates for the age of Gossypieae and the genera in this tribe are much older than previously reported, long-distance oceanic dispersal is still an important process in determining the pantropical distribution of the tribe. The lineages leading to the majority of genera differentiated in the Oligocene and Miocene, when continents were already separated and had attained the main features of their present configuration. The early divergence of *Cienfuegosia c.* 41 Mya (Fig. 6), however, might be associated with vicariance, coincident with the opening of the Drake passage which separated South America from Antarctica–Australia (Scher & Martin, 2006; see below). *Cienfuegosia* is essentially South American with only eight of *c.* 30 species in Africa (Krapovickas, 2003). Its sister clade, which includes the remaining genera of the tribe, is predominantly Palaeotropical–subtropical. The only strictly Neotropical genus in this clade is *Hampea*, which originated much later as result of one dispersal event from Southeast Asia–Oceania to Central America, as indicated by our ancestral range estimation analysis (Fig. 7). *Gossypium*, like *Thespesia*, also has species in America, but these are derived from Old World ancestors (Wendel, Brubaker & Seelanan, 2010). This essentially disjunct distribution of the two deepest sister clades of Gossypieae and a divergence time estimate that coincides with a major geophysical event strongly suggest that vicariance might have played a major role early in the history of the tribe.

Our estimates for the origin and diversification of the eumalvoids and tribes Malveae and Hibisceae are also much older than those reported in previous studies (Koopman & Baum, 2008). Our analysis placed the radiation of this group in the Late Cretaceous, *c.* 75 Mya, whereas the crown groups of Malveae and Hibisceae are at least 55 and 41 Myr old, respectively. Koopman & Baum (2008) dated the radiation of the eumalvoids at *c.* 20 Mya and of Malveae and Hibisceae at 11–14 and 15–19 Mya, respectively, age estimates that are notably young for these diverse, worldwide-distributed groups (the eumalvoids comprise *c.* 1800 species in 111 genera, of which c. 1040 belong to Malveae and c. 630 to Hibisceae). Their analysis relied on two fossil calibrations, the oldest of which, pollen of *Echiperiporites estelae* from the mid- to late Eocene of Venezuela, was used to fix the minimum age of core Malvoideae at 40 Mya.

Echiperiporites estelae is often mentioned in the literature of Malvaceae as the oldest fossil pollen for Malvoideae (Wendel & Albert, 1992; Seelanan *et al.*, 1997; Pfeil *et al.*, 2002; Baum *et al.*, 2004; Koopman & Baum, 2008; Carvalho *et al.*, 2011), despite there being other much older published records (e.g. *Malvacipollis* spp. from the late Palaeocene of North America, Gaponoff, 1984). It has, moreover, palynological features that are regarded as derived in Malvoideae, such as large size, spheroidal shape and large number of apertures (Germeraad, Hopping & Muller, 1968). In fact, *E. estelae* pollen is similar to that of *Hibiscus*, which would place it in the Hibisceae clade of our phylogenetic tree, rather than at the base of core Malvoideae. As noted by Pfeil *et al.* (2002), early-diverging genera of core Malvoideae (*Radyera*, *Camptostemon* Mast. and *Pentaplaris* L.O.Williams & Standl.) have suboblate pollen grains with few (3–10), equatorially distributed apertures, characteristics that are indeed considered ancestral in this group (Christensen, 1986). The different placement of *E. estelae* in our analysis and the inclusion of a higher number of fossils, several of which are much older than *E. estelae* (Appendix 2), resulted in more realistic dates for the origin and radiation of eumalvoids and the constituent tribes.

Based on the current distribution and habitat preference of the early-branching genera of core Malvoideae, and the relatively recent palynological fossil evidence, Baum *et al.* (2004) suggested that the eumalvoids originated in Australasia and that oceanic dispersal played the most important role early in the history of the group. Our much older dates for the origin and radiation of the eumalvoids however, indicate that vicariance may have also contributed to the worldwide distribution and diversification of the group. During the late Cretaceous, land connections existed between high-latitude land masses, notably between Australia and Antarctica and between Antarctica and South America (Scotese, 2004; Reguero *et al.*, 2013). The much warmer climate at the time may have favoured the migration of early malvoids across these corridors, contributing in this way to their range expansion. The complete separation of Antarctica, first from South America by the mid- to late Eocene, *c.* 45–40 Mya (Scotese, 2004; Scher & Martin, 2006), and then from Australia by the late Eocene to early Oligocene, *c.* 35–33 Mya (Stickley *et al.*, 2004), would have isolated the populations which subsequently differentiated into several lineages that radiated worldwide. Future analyses that include more species of eumalvoids will certainly shed further light on the processes behind the diversification of this group.

Taxonomic treatment

Azanza Alef. in Bot. Zeitung 19: 298. 1861.

Type: *Azanza lampas* (Cav.) Alef. (*Hibiscus lampas* Cav.)

Erect shrubs up to 3.5 m, with a sparse to dense vestiture of stellate hairs. Leaves petiolate; stipules lanceolate to subulate, caducous; blade ovate to suborbicular, cordate to subtruncate at base, entire or three-palmately-lobed, with an abaxial nectary on the midrib. Flowers showy, solitary in the axils or in pauciflorous sympodial inflorescences. Pedicel surmounted by three rounded nectaries. Epicalyx of six to nine triangular to filiform, inconspicuous bracteoles, caducous or sometimes persistent. Calyx broadly cupuliform, five-dentate to five-lobed, persistent, the teeth triangular to subulate, up to 15 mm, trinerved. Corolla bright yellow, with a conspicuous maroon centre. Staminal column included, antheriferous throughout. Carpels (three) four or five; style undivided, the four or five stigmatic lobes decurrent. Capsule (tri-) tetra- or pentalocular, dehiscent, ovoid-angled, woody. Seeds 8–12 per locule, small (4–6 mm), glabrous or with sparse, appresed short hairs.

Azanza lampas (Cav.) Alef. in Bot. Zeitung 19: 298. 1861 ≡ *Hibiscus lampas* Cav., Diss. 3: 154 (t. 56, f. 2). 1787 ≡ *Thespesia lampas* (Cav.) Dalzell ex Dalzell & A.Gibson, Bombay Fl. 19. 1861 ≡ *Bupariti lampas* (Cav.) Rothm. in Feddes Repert Spec. Nov. Regni Veg. 53: 7. 1944. TYPE: PHILIPPINES, '*Hibiscus lampas* c. no. 46', 7 Mar 1786, *Sonnerat s.n.* [holotype: P-JU #12356A (photograph!, IDC microfiche #915-A3); holotype fragment: NY (photograph!)].

= *Hibiscus tetralocularis* Roxb., Hort. Beng. 97. 1814, *nom. inval.*, *nom. nud.*; Fl. Ind. 3: 198. 1832. TYPE: INDIA, 'Coromandel', *Roxburgh s.n.* [lectotype, designated by van Borssum Waalkes, 1966: 116: K (n.v.)].

= *Hibiscus callosus* Blume, Bijdr. Fl. Ned. Ind. 2: 67. 1825. TYPE: INDONESIA, 'Java', *Blume s.n.* [lectotype, designated by van Borssum Waalkes, 1966: 116: P (photograph!); isolectotypes: L (×2, photographs)].

= *Pariti gangeticum* G.Don, Gen. Hist. 1: 485. 1831, '*Paritium*' ≡ *Hibiscus gangeticus* Roxb. ex Wight & Arn., Prodr. Fl. Ind. Orient. 1: 49. 1834*, nom. inval.*, *pro syn.* (non *Hibiscus gangeticus* Willd. 1814). TYPE: INDIA, '*Hibiscus gangeticus*', *Roxburgh s.n.* [lectotype, here designated: K (photograph!); isolectotype: BR (photograph!)].

= *Thespesia sublobata* Blanco, Fl. Filip., ed. 2, 382. 1845. TYPE: PHILIPPINES, 'Luzon, Prov. of Rizal, Antipolo', 13 Oct 1913, *Merrill: Spec. Blanc. 561* [neotype, designated by van Borssum Waalkes, 1966: 116: GH!; isoneotypes: BO (n.v), K (photograph!), L (n.v.), MO!, NY!, P (×2, photographs!), US!].

= *Azanza zollingeri* Alef. in Bot. Zeitung 19: 298. 1861 ≡ *Abelmoschus zollingeri* (Alef.) Müll.Berol. in Ann. Bot. Syst. 7: 407. 1868. TYPE: INDONESIA, 'Java', *Zollinger 1203* [lectotype, here designated: BM (photograph!); isolectotype: P (photograph!)].
= *Azanza acuminata* Alef. in Bot. Zeitung 19: 299. 1861 ≡ *Abelmoschus acuminatus* (Alef.) Müll.Berol. in Ann. Bot. Syst. 7: 407. 1868. TYPE: INDIA, 'Terra Canara, prope urbem Mangalor', 1849 [1847], *Hohenacker 388* [lectotype, here designated: BM (photograph!); isolectotypes: L (n.v.), P (photograph!)].
= *Thespesia lampas* var. *longisepala* Borss.Waalk. in Blumea 14: 118. 1966. TYPE: INDONESIA, SE Borneo, Martapura, *Ramali 1930 = bb. 629* [holotype: BO (n.v.)].

Azanza thespesioides (Benth.) F.Areces, ***comb. nov.*** ≡ *Fugosia thespesioides* Benth., Fl. Austral. 1: 220. 1863 ≡ *Hibiscus thespesioides* R.Br ex Benth., Fl. Austral. 1: 220. 1863, *nom. inval.*, *pro syn.* ≡ *Gossypium thespesioides* (Benth.) F.Muell., Fragm. Phyt. Austral. 9: 127. 1875 ≡ *Cienfuegosia thespesioides* (Benth.) Hochr., Ann. Cons. Jard. Bot. Genève 6: 58. 1902 ≡ *Notoxylinon thespesioides* (Benth.) Lewton in J. Wash. Acad. Sci. 5: 306. 1915 ≡ *Thespesia lampas* var. *thespesioides* (Benth.) Fryxell in Aust. J. Bot. 13: 97. 1965 ≡ *Thespesia thespesioides* (Benth.) Fryxell, Nat. Hist. Cotton Tribe: 99. 1979. TYPE: AUSTRALIA, 'N Coast, Island Z', 24 Feb 1803, *Brown 5139* [lectotype, designated by Fryxell 1965: 97: BM (photograph!); isolectotypes: BM (photograph!), BRI (n.v.), CANB (n.v.), K (×2, photographs!), MEL (n.v.), P (photograph!), PERTH (n.v.)].
= *Fugosia flaviflora* F.Muell., Fragm. Phyt. Austral. 5: 44. 1865 ≡ *Gossypium flaviflorum* (F.Muell.) Todaro, Relaz. 105. 1877 ≡ *Hibiscus flaviflorus* (F.Muell.) Kuntze, Revis. Gen. Pl. 1: 69. 1891 ≡ *Cienfuegosia flaviflora* (F.Muell.) Hochr., Ann. Cons. Jard. Bot. Genève 6: 56. 1902 ≡ *Notoxylinon flaviflorum* (F.Muell.) Lewton in J. Wash. Acad. Sci. 5: 307. 1915. TYPE: WESTERN AUSTRALIA, 'Glenelg River', *Martin 6* [holotype: MEL (n.v.)].

Thespesia Sol. ex Corrêa in Ann. Mus. Natl. Hist. Nat. 9: 290. 1807, *nom cons.* ≡ *Bupariti* Duhamel, Semis Plantat. Arbres, Add.: 5. 1760, *nom rej.* ≡ *Pariti* Adans., Fam. Pl. 2: 401. 1763, *nom. illeg.*

Type: *Thespesia populnea* (L.) Sol. ex Corrêa [*Hibiscus populneus* L., *Bupariti populnea* (L.) Rothm.].
= *Montezuma* DC., Prodr. 1: 477. 1824. Type: *Montezuma speciosissima* DC. (= *Thespesia grandiflora* DC.).
= *Maga* Urb., Symb Antill. 7: 281. 1912. Type: *Maga grandiflora* (DC.) Urb. (*Thespesia grandiflora* DC.).
= *Ulbrichia* Urb. in Dansk Bot. Ark. 4 (7): 7. 1924. Type: *Ulbrichia beatensis* Urb. [*Thespesia beatensis* (Urb.) Fryxell].
= *Shantzia* Lewton in J. Wash. Acad. Sci. 18: 15. 1928. Type: *Shantzia garckeana* (F.Hoffm.) Lewton (*Thespesia garckeana* F.Hoffm.).
= *Armouria* Lewton in J. Wash. Acad. Sci. 23: 64. 1933. Type: *Armouria beata* Lewton [*Thespesia beatensis* (Urb.) Fryxell].
= *Atkinsia* R.A.Howard in Bull. Torrey Bot. Club 76: 97. 1949. Type: *Atkinsia cubensis* (Britton & P.Wilson) R.A.Howard [*Maga cubensis* Britton & P.Wilson, *Thespesia cubensis* (Britton & P.Wilson) J.B.Hutch.].
= *Thespesiopsis* Exell & Hillc. in Mendonca, Contrib. Conhec. Fl. Mocamb. 2 (Estud., Ens. & Docum. 12): 55. 1954. Type: *Thespesiopsis mossambicensis* Exell & Hillc. [*Thespesia mossambicensis* (Exell & Hillc.) Fryxell].

Tall shrubs or small trees to trees up to 30 m, with a vestiture of scales or stellate hairs. Leaves petiolate; stipules lanceolate to subulate, caducous; blade ovate to deltoid, cordate to subtruncate at base, entire or less often three- to five-palmately-lobed, usually with an abaxial nectary on the midrib, sometimes with domatia at the base. Flowers showy, solitary in the axils, sometimes grouped terminally by reduction of the internodes. Pedicel often articulate, erect or occasionally drooping. Epicalyx of 3–27 linear to lanceolate bracteoles, persistent or caducous. Calyx cupuliform, truncate, entire to five-denticulate (the teeth rarely subulate and long), persistent (rarely caducous). Corolla yellow, less often white or deep rose, usually with a dark red centre. Staminal column included (rarely exserted), antheriferous throughout or in distal half. Carpels (four) five; style undivided, the four or five stigmatic lobes decurrent. Capsule (tetra-) pentalocular, indehiscent or rarely dehiscent, oblate to globose or less commonly ovoid, woody to coriaceous or sometimes fleshy. Seeds one to four per locule, large (7–15 mm), glabrous to densely pubescent.

ACKNOWLEDGEMENTS

We thank the following people and institutions for providing leaf samples: Pedro A. González (Centro de Investigaciones y Servicios Ambientales y Tecnológicos, Cuba), Banesa Falcón and Rosalina Berazain (HAJB), Marcos Caraballo (Pennsylvania State University), Rachun Pooma (BKF), the Nationaal Herbarium Nederland (L), The Australian National Herbarium (CANB) and The National Herbarium of New South Wales (NSW). The DNA Bank at Kew

provided genomic DNA. We also thank the curators of AAU, B, BM, F, GH, ISC, K, L, MO, NY and US for the loan of specimens and/or supplying images for study and of GH, MO, NH and NU for permission to sample material from specimens. Thanks are extended to the Ministerio de Medio Ambiente y Recursos Naturales of Dominican Republic for research permits, and Amelia Mateo and Gerson Feliz (Grupo Jaragua) for enormously facilitating our fieldwork in Beata Island. Víctor Vega, Kristopher Bonefont and Jaileen Merced (University of Puerto Rico) provided fieldwork assistance in Puerto Rico and Dominican Republic. Danny Rojas (University of Aveiro, Portugal) and Angel Vale (University of Vigo, Spain) provided invaluable help with biogeographical analyses. Pedro A. González, Craig Gibbon, Dinesh Valke, Alex Monro, Wilnelia Recart and Stephan Dressler kindly supplied or gave permission to use photographs. We are also grateful to the staff of the Evolutionary Biology lab at the University of Puerto Rico-Rio Piedras, and especially to Yadira Ortiz for assistance with laboratory work. This study was supported in part by the Center for Applied Tropical Ecology and Conservation of the University of Puerto Rico (NSF HRD-0734826). F.A.B. also thanks the Dean of Graduate Studies and Research and the Department of Biology of the University of Puerto Rico in Río Piedras for financial support and Professor Elvia Meléndez for the invitation to attend the South African PhD Project conference in 2012 during which it was possible to visit South African herbaria. Sequencing was conducted at the Sequencing and Genomics Facility of the UPR in Río Piedras and MSRC/UPR, which is funded by NIH/NIGMS-Award Number P20GM103475. We specifically acknowledge the assistance of technicians Silvia Planas and Dania Rodríguez. This work is part of F.A.B.'s PhD dissertation at the University of Puerto Rico.

REFERENCES

Acevedo-Rodríguez P, Strong M. 2008. Floristic richness and affinities in the West Indies. *Botanical Review* **74:** 5–36.

Alain H. 1958. La flora de Cuba: sus principales características y su origen probable. *Revista de la Sociedad Cubana de Botánica* **15:** 84–96.

Alefeld FGC. 1861. Über *Hibiscus lampas* Cav. *Botanische Zeitung* **19:** 297–299.

Andrus N, Tye A, Nesom G, Bogler D, Lewis C, Noyes R, Jaramillo P, Francisco-Ortega J. 2009. Phylogenetics of *Darwiniothamnus* (Asteraceae: Astereae) – molecular evidence for multiple origins in the endemic flora of the Galápagos Islands. *Journal of Biogeography* **36:** 1055–1069.

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

Barreda VD. 1993. Late Oligocene?–Miocene pollen of the families Compositae, Malvaceae and Polygonaceae from the Chenque Formation, Golfo San Jorge Basin, southeastern Argentina. *Palynology* **17:** 169–186.

Bates DM. 1990. Malvaceae. In: Wagner WL, Herbst DR, Sohmer SH, eds. *Manual of the flowering plants of Hawai'i, Vol. 1. Bishop Museum Special Publications 83*. Honolulu: University of Hawaii and Bishop Museum Press, 868–903.

Baum DA, DeWitt Smith S, Yen A, Alverson WS, Nyffeler R, Whitlock BA, Oldham RL. 2004. Phylogenetic relationships of *Malvatheca* (Bombacoideae and Malvoideae; Malvaceae *sensu lato*) as inferred from plastid DNA sequences. *American Journal of Botany* **91:** 1863–1871.

Beauchamp J, Lemoigne Y. 1973. Description d'une paleoflora du Crétacé terminal Eocéne dans le massif du Chercher (Province d' Harar, Ethiopie). *Documents des Laboratories de Géologie de la Faculté des Sciences de Lyon* **56:** 167–180.

Bell AA, Stipanovic RD, Howell CR, Fryxell PA. 1975. Antimicrobial terpenoids of *Gossypium*: hemigossypol 6-methoxyhemigossypol, and 6-deoxyhemigossypol. *Phytochemistry* **14:** 225–231.

Borhidi A. 1991. *Phytogeography and vegetation ecology of Cuba*. Budapest: Akadémiai Kiadó.

van Borssum Waalkes J. 1966. Malesian Malvaceae revised. *Blumea* **14:** 1–213.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10:** e1003537.

Burton KW, Ling H-F, O'Nions RK. 1997. Closure of the Central American Isthmus and its effect on deep-water formation in the North Atlantic. *Nature* **386:** 382–385.

Carvalho MR, Herrera FA, Jaramillo CA, Wing SL, Callejas R. 2011. Paleocene Malvaceae from northern South America and their biogeographical implications. *American Journal of Botany* **98:** 1337–1355.

Christensen PB. 1986. Pollen morphological studies in the Malvaceae. *Grana* **25:** 95–117.

Chumbhale DS, Upasani CD. 2012a. Hepatoprotective and antioxidant activity of *Thespesia lampas* (Cav.) Dalz & Gibs. *Phytopharmacology* **2:** 114–122.

Chumbhale DS, Upasani CD. 2012b. Pharmacognostic standardization of stems of *Thespesia lampas* (Cav.) Dalz & Gibs. *Asian Pacific Journal of Tropical Biomedicine* **2:** 357–363.

Cronn RC, Small RL, Haselkorn T, Wendel JF. 2002. Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. *American Journal of Botany* **89:** 707–725.

Cronn RC, Small RL, Wendel JE. 1999. Duplicated genes evolve independently following polyploid formation in cotton. *Proceedings of the National Academy of Sciences USA* **96:** 14406–14411.

Dhanarasu S, Al-hazimi AM, Sethuraman P, Selvam M. 2010. Chemopreventive and antilipidperoxidative potential

of *Thespesia populnea* (L.) on experimental buccal pouch carcinogenesis. *Ibnosina Journal of Medicine and Biomedical Sciences* **2:** 269–277.

Drábková L, Kirschner J, Vlček Č. 2002. Comparison of seven DNA extraction and amplification protocols in historical herbarium specimens of Juncaceae. *Plant Molecular Biology Reporter* **20:** 161–175.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 214.

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

Duque-Caro J. 1990. Neogene stratigraphy, paleoceanography and paleobiogeography in northwest South America and the evolution of the Panama Seaway. *Palaeogeography, Palaeoclimatology, Palaeoecology* **77:** 203–234.

Escobar García P, Schönswetter P, Fuertes Aguilar J, Nieto Feliner G, Schneeweiss GM. 2009. Five molecular markers reveal extensive morphological homoplasy and reticulate evolution in the *Malva* alliance (Malvaceae). *Molecular Phylogenetics and Evolution* **50:** 226–239.

Exell AW, Hillcoat D. 1954. Um novo género de Malvaceae de Moçambique. *Contribuições para o conhecimento da flora de Moçambique* **2:** 55–61.

Francis JK. 1999. Especies forestales para plantar en áreas forestales, rurales y urbanas de Puerto Rico. USDA Forest Service, General Technical Report IITF-13.

Fredericksen NO. 1980. Paleogene sporomorphs from South Carolina and quantitative correlations with the Gulf Coast. *Palynology* **4:** 125–179.

Friday JB, Okano D. 2006. *Thespesia populnea* (milo). In: Elevitch CR, ed. *Traditional trees of Pacific islands: their culture, environment and use*. Holualoa: Permanent Agriculture Resources, 757–774.

Fryxell PA. 1968. A redefinition of the tribe Gossypieae. *Botanical Gazette* **129:** 296–308.

Fryxell PA. 1969. The genus *Hampea* (Malvaceae). *Brittonia* **21:** 359–396.

Fryxell PA. 1979. *The natural history of the cotton tribe*. College Station, TX: Texas A&M University Press.

Fryxell PA. 1992. A revised taxonomic interpretation of *Gossypium* L. (Malvaceae). *Rheedea* **2:** 108–165.

Fryxell PA. 1997. The American genera of Malvaceae –II. *Brittonia* **49:** 204–269.

Fuertes Aguilar J, Fryxell PA, Jansen RK. 2003. Phylogenetic relationships and classification of the *Sida* generic alliance based on nrDNA ITS evidence. *Systematic Botany* **28:** 352–364.

Gaponoff SL. 1984. Palynology of the Silverado Formation (Late Paleocene), Riverside and Orange Counties, California. *Palynology* **8:** 71–106.

Germeraad JH, Hopping CA, Muller J. 1968. Palynology of Tertiary sediments from tropical areas. *Review of Palaeobotany and Palynology* **6:** 189–348.

Graham A. 1976. Studies in Neotropical paleobotany. II. The Miocene communities of Veracruz, Mexico. *Annals of the Missouri Botanical Garden* **63:** 787–842.

Graham A. 1988. Studies in Neotropical paleobotany. V. The Lower Miocene communities of Panama – the Culebra Formation. *Annals of the Missouri Botanical Garden* **75:** 1440–1466.

Graham A. 1991. Studies in Neotropical paleobotany. IX. The Pliocene communities of Panama – angiosperms (dicots). *Annals of the Missouri Botanical Garden* **78:** 201–223.

Graham A. 1994. Neogene palynofloras and terrestrial paleoenvironments in northern Latin America. In: Thompson RS, ed. *Pliocene terrestrial environments and data/model comparisons*, abstracts from a USGS workshop, Herndon, Virginia, May 22–23, 1993. U.S. Department of the Interior, U.S. Geological Survey, Open-File Report 94–23, 23–30.

Graham A. 2003. Historical phytogeography of the Greater Antilles. *Brittonia* **55:** 357–383.

Graham A, Cozadd D, Areces-Mallea A, Frederiksen NO. 2000. Studies in Neotropical paleobotany. XIV. A palynoflora from the middle Eocene Saramaguacán Formation of Cuba. *American Journal of Botany* **87:** 1526–1539.

Graham A, Jarzen DM. 1969. Studies in Neotropical paleobotany. I. The Oligocene communities in Puerto Rico. *Annals of the Missouri Botanical Garden* **56:** 306–357.

Graham SA. 2010. Revision of the Caribbean genus *Ginoria* (Lythraceae), including *Haitia* from Hispaniola. *Annals of the Missouri Botanical Garden* **97:** 34–90.

Grimm GW, Kapli P, Bomfleur B, Mcloughlin S, Renner SS. 2015. Using more than the oldest fossils: dating Osmundaceae with three Bayesian clock approaches. *Systematic Biology* **64:** 396–405.

Gros JP. 1992. Les bois fossiles d'Ethiopie. *Inventaire systematique. Geobios* **25:** 3–12.

Gyory J, Mariano AJ, Ryan EH. 2016. The Caribbean current, ocean surface currents. Available at: http://oceancurrents.rsmas.miami.edu/caribbean/caribbean.html (accessed 1 February 2016).

Heath TA, Huelsenbeck JP, Stadler T. 2014. The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences USA* **111:** E2957–E2966.

Hollick A. 1928. *Paleobotany of Puerto Rico. Scientific survey of Porto Rico and the Virgin Islands, Vol. 7*, part 3. New York: New York Academy of Sciences.

Hollick A. 1936. The Tertiary floras of Alaska. *US Geological Survey Professional Paper* **182:** 1–185.

Howard RA. 1949. *Atkinsia*, gen. nov., *Thespesia* and related West Indian genera of the Malvaceae. *Bulletin of the Torrey Botanical Club* **76:** 89–100.

Howard RA. 1973. The vegetation of the Antilles. In: Graham A, ed. *Vegetation and vegetational history of northern Latin America*. Amsterdam: Elsevier Scientific, 1–38.

Iturralde-Vinent MA. 2006. Meso-Cenozoic Caribbean paleogeography: implications for the historical biogeography of the region. *International Geology Review* **48:** 791–827.

Iturralde-Vinent MA, MacPhee R. 1999. Paleogeography of the Caribbean region, implications for Cenozoic biogeog-

raphy. *Bulletin of the American Museum of Natural History* **238:** 1–95.

Jaramillo C, Moreno E, Ramírez V, da Silva S, de la Barrera A, de la Barrera A, Sánchez C, Morón S, Herrera F, Escobar J, Koll R, Manchester SR, Hoyos N. 2014. Palynological record of the last 20 Million years in Panama. In: Stevens WD, Montiel OM, Raven P, eds. *Paleobotany and biogeography: a festschrift for Alan Graham in his 80th year*. St Louis: Missouri Botanical Garden Press, 134–253.

Johnson LA, Soltis DE. 1994. *matK* DNA sequences and phylogenetic reconstruction in Saxifragaceae *s. str*. *Systematic Botany* **19:** 143–156.

Jordan WC, Courtney MW, Neigel JE. 1996. Low levels of intraspecific genetic variation at a rapidly evolving chloroplast DNA locus in North American duckweeds (Lemnaceae). *American Journal of Botany* **83:** 430–439.

Keigwin LD. 1978. Pliocene closing of the Isthmus of Panama, based on biostratigraphic evidence from nearby Pacific Ocean and Caribbean Sea cores. *Geology* **6:** 630–634.

Keigwin LD. 1982. Isotope paleoceanography of the Caribbean and east Pacific: role of Panama uplift in late Neogene time. *Science* **217:** 350–353.

Kelchner SA, Clark LG. 1997. Molecular evolution and phylogenetic utility of the chloroplast *rpl16* intron in *Chusquea* and the Bambosoideae (Poaceae). *Molecular Phylogenetics and Evolution* **8:** 385–397.

Koopman MM, Baum DA. 2008. Phylogeny and biogeography of tribe Hibisceae (Malvaceae) on Madagascar. *Systematic Botany* **33:** 364–374.

Kosalge SB, Fursule RA. 2009. Investigation of *in vitro* anthelmintic activity of *Thespesia lampas* (Cav.). *Asian Journal of Pharmaceutical and Clinical Research* **2:** 69–71.

Krapovickas A. 2003. Las especies austroamericanas del género *Cienfuegosia* Cav. (Malvaceae Gossypieae). *Bonplandia* **12:** 5–47.

Kräusel R. 1939. Ergebnisse der Forschungsreisen Prof. E. Stromers in den Wüsten Ägyptens, IV. Die fossilen Floren Ägyptens: 3. Die fossilen Planzen Ägyptens. E–L. *Abhandlungen Bayerischen Akademie der Wissenschaften, Mathematisch-naturwissenschaftliche Abteilung (N.F.)* **47:** 1–140.

La Duke JC, Doebley J. 1995. A chloroplast DNA based phylogeny of the Malvaceae. *Systematic Botany* **20:** 259–271.

Landis MJ, Matzke NJ, Moore BR, Huelsenbeck JP. 2013. Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology* **62:** 789–804.

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

Leigh EG, O'Dea A, Vermeij GJ. 2014. Historical biogeography of the Isthmus of Panama. *Biological Reviews of the Cambridge Philosophical Society* **89:** 148–172.

Lewton FL. 1928. *Shantzia*, a new genus of African shrubs related to *Gossypium*. *Journal of the Washington Academy of Science* **18:** 10–16.

Little EL, Wadsworth FH. 1964. *Common trees of Puerto Rico and the Virgin Islands. Agriculture Handbook 249*. Washington, DC, USDA Forest Service.

Liu K, Warnow T. 2014. Large-scale multiple sequence alignment and tree estimation using SATe. *Methods in Molecular Biology* **1079:** 219–244.

Liu K, Warnow TJ, Holder MT, Nelesen SM, Yu J, Stamatakis AP, Linder CR. 2012. SATe-II: very fast and accurate simultaneous estimation of multiple sequence alignments and phylogenetic trees. *Systematic Biology* **61:** 90–106.

Lykke AM. 1994. The vegetation of Delta du Saloum National Park, Senegal. AAU Reports 33, Department of Systematic Botany, Aarhus University.

MacPhail MK, Truswell EM. 1989. Palynostratigraphy of the central west Murray Basin. *BMR Journal of Australian Geology and Geophysics* **11:** 301–331.

Maddison WP, Maddison DR. 2011. Mesquite: a modular system for evolutionary analysis, v. 2.75. Available at: http://mesquiteproject.org (accessed 7 December 2013).

Matzke NJ. 2013a. BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R scripts. R package, v. 0.2.1. Available at: http://CRAN.R-project.org/package=BioGeoBEARS (accessed 13 August 2015).

Matzke NJ. 2013b. Probabilistic historical biogeography: new models for founder event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography* **5:** 242–248.

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

Matzke NJ. 2014b. Stochastic mapping under biogeographical models. PhyloWiki BioGeoBEARS website. Available at: http://phylo.wikidot.com/biogeobears#stochastic_mapping (accessed 13 August 2015).

Maunder M, Abdo M, Berazain R, Clubbe C, Jiménez F, Leiva A, Santiago-Valentín E, Francisco-Ortega J. 2011. The plants of the Caribbean islands: a review of the biogeography, diversity and conservation of a storm-battered biodiversity hotspot. In: Bramwell D, Caujapé-Castells J, eds. *The biology of island floras*. London: Cambridge University Press, 154–178.

Mautino LR, Cuadrado GA, Anzótegui LM. 2004. Novedades taxonómicas, diversidad y significado evolutivo del polen de Malvaceae en el Terciario de Argentina. *Revista Española de Micropaleontología* **36:** 467–483.

Mojeremane W, Tshwenyane SO. 2004. *Azanza garckeana*: a valuable edible indigenous fruit tree of Botswana. *Pakistan Journal of Nutrition* **3:** 264–267.

Montes C, Cardona A, Jaramillo C, Pardo A, Silva JC, Valencia V, Ayala C, Pérez-Angel LC, Rodriguez-Parra LA, Ramirez V, Niño H. 2015. Middle Miocene closure of the Central America Seaway. *Science* **348:** 226–249.

Mort ME, Randle CP, Kimball RT, Tadesse M, Crawford DJ. 2008. Phylogeny of Coreopsidae (Asteraceae) inferred from nuclear and plastid DNA sequences. *Taxon* **57:** 109–120.

Motley TJ, Wurdack KJ, Delprete PG. 2005. Molecular systematics of the Catesbaeeae–Chiococceae complex (Rubi-

aceae): flower and fruit evolution and biogeographic implications. *American Journal of Botany* **92:** 316–329.

Nagy E. 1962. New pollen species from the Lower Miocene of the Bakony Mountain (Várpalota) of Hungary. *Acta Botanica Academiae Scientiarum Hungaricae* **8:** 153–170.

Namoff S, Luke Q, Jiménez F, Veloz A, Lewis CE, Sosa V, Maunder M, Francisco-Ortega J. 2010. Phylogenetic analyses of nucleotide sequences confirm a unique plant intercontinental disjunction between tropical Africa, the Caribbean, and the Hawaiian Islands. *Journal of Plant Research* **123:** 57–65.

Oleas N, Jestrow B, Calonje M, Peguero B, Jiménez F, Rodríguez-Peña R, Oviedo R, Santiago-Valentín E, Meerow AW, Abdo M, Maunder M, Griffith MP, Francisco-Ortega J. 2013. Molecular systematics of threatened seed plant species endemic in the Caribbean islands. *The Botanical Review* **79:** 528–541.

Olmstead RG, Sweere JA. 1994. Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. *Systematic Biology* **43:** 467–481.

Palmer JD. 1991. Plastid chromosomes: structure and evolution. In: Bogorad L, Vasil IK, eds. *The molecular biology of plastids, cell culture and somatic cell genetics of plants*, Vol. **7A**. San Diego**:** Academic Press, 5–53.

Parthasarathy R, Ilavarasan R, Karrunakaran CM. 2009. Antidiabetic activity of *Thespesia populnea* bark and leaf extract against streptozotocin induced diabetic rats. *International Journal of PharmTech Research* **1:** 1069–1072.

Pfeil BE, Brubaker CL, Craven LA, Crisp MD. 2002. Phylogeny of *Hibiscus* and the tribe Hibisceae (Malvaceae) using chloroplast DNA sequences of *ndhF* and the *rpl16* intron. *Systematic Botany* **27:** 333–350.

Phuphathanaphong L. 2006. *Thepparatia* (Malvaceae), a new genus from Thailand. *Thai Forest Bulletin (Botany)* **34:** 195–200.

Pratap Chandran R, Manju S, Vysakhi MV, Shaji PK, Achuthan NG. 2014. Antibacterial and antifungal activities of *Thespesia populnea* leaf extracts against human pathogens. *International Journal of PharmTech Research* **6:** 290–297.

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

Rambaut A, Drummond AJ. 2014a. LogCombiner v. 2.3.0. Part of the BEAST package. Available at: http://beast2.org/ (accessed 27 June 2015).

Rambaut A, Drummond AJ. 2014b. TreeAnnotator v. 2.3.0. Part of the BEAST package. Available at: http://beast2.org/ (accessed 27 June 2015).

Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Tracer v. 1.6.0. Available at: http://beast.bio.ed.ac.uk/Tracer (accessed 17 January 2015).

Reguero M, Goin F, Acosta HC, Dutra T, Marenssi S. 2013. *Late Cretaceous/Paleogene west Antarctica terrestrial biota and its intercontinental affinities*. Berlin: Springer.

Revell LJ. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3:** 217–223.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61:** 539–542.

Sánchez Andrade A. 2005. *Filogenética molecular de los Espeletiinae, una radiación adaptativa andina*. Tésis de Maestria: Universidad de los Andes, Colombia.

Santiago-Valentin E, Olmstead RG. 2004. Historical biogeography of Caribbean plants: introduction to current knowledge and possibilities from a phylogenetic perspective. *Taxon* **53:** 299–319.

Scher HD, Martin EE. 2006. The timing and climatic influence of the opening of Drake passage. *Science* **312:** 428–430.

Schneider JV, Schulte K, Fuertes Aguilar J, Huertas ML. 2011. Molecular evidence for hybridization and introgression in the Neotropical coastal desert-endemic *Palaua* (Malveae, Malvaceae). *Molecular Phylogenetics and Evolution* **60:** 373–384.

Scotese CR. 2004. Cenozoic and Mesozoic paleogeography: changing terrestrial biogeographic pathways. In: Lomolino MV, Heaney LR, eds. *Frontiers of biogeography: new directions in the geography of nature*. Sunderland, MA: Sinauer Associates, 9–26.

Seelanan T, Schnabel A, Wendel JF. 1997. Congruence and consensus in the cotton tribe (Malvaceae). *Systematic Botany* **22:** 259–290.

Selva Kumar S, Ram Krishna Rao M, Pavunraj M, Bhattacharya D. 2012. Antibacterial therapeutic role of *Thespesia populnea*, *Nerium odorum* and *Ocimum basilicum* on selected human pathogenic microorganisms. *Journal of Pharmacy Research* **5:** 4271–4274.

Shrivastav S, Sindhu RK, Kumar S, Kumar P. 2009. Anti-psoriatic and phytochemical evaluation of *Thespesia populnea* bark extracts. *International Journal of Pharmacy and Pharmaceutical Sciences* **1**(Suppl. 1)**:** 176–185.

Silvestro D, Michalak I. 2012. RaxmlGUI: a graphical front-end for RAxML. *Organisms Diversity & Evolution* **12:** 335–337.

Small RL. 2004. Phylogeny of *Hibiscus* sect. *Muenchhusia* (Malvaceae) based on chloroplast *rpL16* and *ndhF*, and nuclear ITS and GBSSI sequences. *Systematic Botany* **29:** 385–392.

Stamatakis A. 2006. Raxml-vi-hpc: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22:** 2688–2690.

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

Stickley CE, Brinkhuis H, Schellenberg SA, Sluijs A, Röhl U, Fuller M, Grauert M, Huber M, Warnaar J, Williams GL. 2004. Timing and nature of the deepening of the Tasmanian Gateway. *Paleoceanography* **19:** 1–18.

Tate JA. 2011. The status of *Urocarpidium* (Malvaceae): insight from nuclear and plastid-based phylogenies. *Taxon* **60:** 1330–1338.

Tate JA, Fuertes Aguilar J, Wagstaff SJ, La Duke JC, Bodo Slotta TA, Simpson BB. 2005. Phylogenetic relationships within the tribe Malveae (Malvaceae, subfamily Malvoideae) as inferred from ITS sequence data. *American Journal of Botany* **92:** 584–602.

Thiers B. 2015. Index herbariorum: a global directory of public herbaria and associated staff. The New York Botanical Garden, Available at: http://sweetgum.nybg.org/ih/ (accessed 25 October 2015).

Vasudevan M, Parle M. 2007. Memory-enhancing activity of *Thespesia populnea* in rats. *Pharmaceutical Biology* **45:** 267–273.

Verdcourt B, Mwachala GM. 2009. Malvaceae. In: Beentje HJ, Ghazanfar SA, eds. *Flora of Tropical East Africa*. Kew: Royal Botanic Gardens, 1–169.

Wendel JF, Albert VA. 1992. Phylogenetics of the cotton genus (*Gossypium* L.): character-state weighted parsimony analysis of chloroplast DNA restriction site data and its systematic and biogeographic implications. *Systematic Botany* **17:** 115–143.

Wendel JF, Brubaker CL, Seelanan T. 2010. The origin and evolution of *Gossypium*. In: StewartmJMcD,Oosterhuis, DM, Heitholt JJ, Mauney JR, eds. *Physiology of cotton*. New York: Springer, 1–18.

White TJ, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds. *PCR protocols: a guide to methods and applications*. New York: Academic Press, 315–322.

Wood GR. 1986. Late Oligocene to Early Miocene palynomorphs from GSQ Sandy Cape 1–3R. *Geological Survey of Queensland, Publication* **387:** 1–27.

Woodcock DW, Manchester CA. 1998. Fossil cotton from the Salt Lake Crater area, O'ahu, Hawai'i. *Records of the Hawaii Biological Survey for 1997, Part 2: Notes. Bishop Museum Occasional Papers* **56:** 17–19.

Yang Z, Kumar S, Nei M. 1995. A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics* **141:** 1641–1650.

Zamaloa MdelC. 2000. Palinoflora y ambiente en el Terciario del nordeste de Tierra del Fuego, Argentina. *Revista del Museo Argentino de Ciencias Naturales, n.s.* **2:** 43–51.

Zamaloa MdelC, Romero EJ. 1990. Some spores and pollen from the Cullen Formation (Upper Eocene to middle Oligocene), Tierra del Fuego, Argentina. *Palynology* **14:** 123–133.

Appendix 1 Species included in this study with voucher information and GenBank accession numbers. Underlined accession numbers correspond to previously published sequences. Herbarium acronyms follow Index Herbariorum (Thiers, 2015).

Species	Code	Voucher	Origin	GenBank accession number				
				trnk/matK	*ndhF*	*rpl16*	ITS	*CesA1*
Abutilon abutiloides (Jacq.) Garcke	AA	*F. Areces et al. 580* (UPRRP)	Dominican Republic	KT966998	KT967036	KT967074	KT966960	–
Alcea rosea L.	AlRos			EU346805	EU346847	–	AH010172	–
Alyogyne hakeifolia (Giord.) Alef.	AlHak			AY589059	AY589083	AF384564	–	–
Bastardia bivalvis (Cav.) Kunth	BB	*F. Areces et al. 571* (JBSD, UPRRP)	Dominican Republic	KT966996	KT967034	KT967072	KT966958	–
Cephalohibiscus peekelii Ulbr.	CP1	*J. Regalado and P. Katik 1192* (GH)	Papua New Guinea	KT966964	KT967003	KT967040	KT966926	KT966904
Cienfuegosia heterophylla (Vent.) Garcke	CH	*P. A. González 1202-13* (HAJB, UPRRP)	Cuba	KT966962	KT967001	KT967038	KT966924	KT966902
Cienfuegosia tripartita Gürke	CT			–	U55324	–	U56777	AF376047
Gossypioides kirkii (Mast.) Skovsted	GK			AF403563	U55329	AF403104	U56783	AF201886
Gossypium anomalum Wawra & Peyr.	GAn			AF403557	U55332	AF403100	U56806	AF419970
Gossypium arboreum L.	GAr			HQ325740	U55331	AF031451	U12712	EU626444
Gossypium barbadense L.	GBa			AP009123	U55339	AF031453	U12715	HQ143019
Gossypium bickii Prokh.	GBi			AF403562	AF403555	AF059428	U56787	AF419973
Gossypium hirsutum L.	GHi			HQ901196	U55340	AF031452	U12719	AF139443
Gossypium longicalyx J.B.Hutch. & B.J.S.Lee	GLon			AF403561	U55338	AF403103	U12722	AF419972
Gossypium raimondii Ulbr.	GRa			AF403559	U55335	AF403101	U12718	HQ143022
Gossypium robinsonii F.Muell.	GRo			AF403558	U55334	AF059425	U12710	AF139446
Gossypium stocksii Mast.	GSt			JF317355	U55337	NC_023217	U56812	–
Gossypium turneri Fryxell	GTu			AF520731	U55336	AF520721	U12726	AY125069
Hampea appendiculata (Donn.Sm) Standl.	HAp			–	U55327	–	–	AF376042
	HAp	*D. Santamaría 8869* (INB)	Costa Rica	KT966970	–	KT967046	KT966932	–
Hampea mexicana Fryxell	HMe	*E. Martínez and G. Aguilar 36537* (MO)	Mexico	KT966968	KT967007	KT967044	KT966930	KT966908
Hampea nutricia Fryxell	HNu	*G. Ibarra Manriquez 5373* (MO)	Mexico	KT966967	KT967006	KT967043	KT966929	KT966907
Hampea trilobata Standl.	HTr	*M. J. M. Christenhusz and C. Trejo 6073* (BM)	Mexico	KT966969	KT967008	KT967045	KT966931	KT966909

Table . *Continued*

Species	Code	Voucher	Origin	GenBank accession number				
				trnk/matK	*ndhF*	*rpl16*	ITS	*CesA1*
Hibiscus striatus Cav.	HS	*R. Moran 8723 and F. Areces* (INB)	Costa Rica	KT966990	KT967028	KT967066	KT966952	–
Howittia trilocularis F.Muell.	HoTri			AY589065	AY589085	AF384615	AY591832	–
Kokia drynarioides Lewton	KDr			AF403564	U55330	AF403105	U56784	AF376040
Kokia kauaiensis (Rock) O.Deg	KK			–	–	–	U56785	–
& Duvel	KK	*J. D. Ackerman et al. 4542* (UPRRP)	Hawaii	KT966961	KT967000	KT967037	–	KT966901
Lagunaria patersonia (Andrews) G.Don	LagPat			AY589064	AY589084	AF384616	–	–
Lebronnecia kokioides Fosberg	LK	*J. D. Ackerman et al. 4538* (UPRRP)	Hawaii (cult. from Marquesas Islands)	KT966963	KT967002	KT967039	KT966925	KT966903
Malachra radiata L.	MR	*R. Moran 8724 and F. Areces* (INB)	Costa Rica	KT966993	KT967031	KT967069	KT966955	–
Malope trifida Cav.	MaTri			AY589060	AY589076	–	EF419532	–
Malva parviflora L.	MP			–	EU346844	FJ204762	AF303031	–
	MP	*F. Areces and V. Vega 324* (UPTC)	Colombia	KT966999	–	–	–	–
Malvaviscus penduliflorus DC.	MAr	*F. Areces and V. Vega 925* (UPRRP)	Puerto Rico	KT966994	KT967032	KT967070	KT966956	–
Modiola caroliniana (L.) G.Don	MoCar			EF207256	EF207287	JF799586	AY172190	–
Pavonia fruticosa (Mill.) Fawc. & Rendle	PF	*F. Areces et al. 605* (UPRRP)	Puerto Rico	KT966991	KT967029	KT967067	KT966953	–
Pavonia spinifex (L.) Cav.	PSp	*F. Areces et al. 602* (UPRRP)	Puerto Rico	KT966992	KT967030	KT967068	KT966954	–
Radyera farragei (F.Muell) Fryxell & S.H.Hashmi	RadFar			AY589063.1	AY589078.1	AF384623.1	–	–
Sida rhombifolia L.	SR	*F. Areces and V. Vega 610* (UPRRP)	Puerto Rico	KT966997	KT967035	KT967073	KT966959	–
Sphaeralcea angustifolia (Cav.) G.Don	SphAn			EF207255	EF207286	–	AY172192	–
Thepparatia thailandica Phuph.	TT	*R. Pooma s.n.* (UPRRP)	Thailand	KT966989	KT967027	KT967065	KT966951	–
Thespesia acutiloba (Baker f.) Exell & Mendonça	TA1	*D. Styles 2689* (NH)	South Africa	KT966978	KT967016	KT967054	KT966940	KT966916

Table . *Continued*

Species	Code	Voucher	Origin	GenBank accession number				
				trnk/matK	*ndhF*	*rpl16*	ITS	*CesA1*
Thespesia beatensis (Urb.) Fryxell	TB4	*F. Areces et al. 577* (JBSD, UPRRP)	Dominican Republic	KT966971	KT967009	KT967047	KT966933	KT966910
	TB12	*F. Areces et al. 582* (UPRRP)	Dominican Republic	KT966972	KT967010	KT967048	KT966934	KT966911
Thespesia cubensis (Britton & P.Wilson) J.B.Hutch.	TC5	*P. A. Gonzalez 1202-6* (HAJB, UPRRP)	Cuba	KT966975	KT967013	KT967051	KT966937	KT966914
	TC18	*R. Berazain et al. HFC 87787* (HAJB)	Cuba	KT966976	KT967014	KT967052	KT966938	KT966915
Thespesia danis Oliv.	TD2	*M. A. Mwangoka and A. Maingo 1333* (MO)	Tanzania	KT966977	KT967015	KT967053	KT966939	–
Thespesia fissicalyx Borss.Waalk.	TF7	*W. Takeuchi et al. 13739* (GH)	Papua New Guinea	KT966988	KT967026	KT967064	KT966950	–
Thespesia garckeana F.Hoffm.	TGa1	*N. A. Mwangulango 1153* (MO)	Tanzania	KT966979	KT967017	KT967055	KT966941	KT966917
	TGa2	*O. A. Kibure 668* (MO)	Tanzania	KT966980	KT967018	KT967056	KT966942	KT966918
Thespesia gummiflua Capuron	TGu3	*R. Ramananjanahary et al. 31* (MO)	Madagascar	KT966981	KT967019	KT967057	KT966943	–
Thespesia grandiflora DC.	TG5	*F. Areces et al. 604* (UPRRP)	Puerto Rico	KT966973	KT967011	KT967049	KT966935	KT966912
	TG30	*F. Areces and V. Vega 761* (UPRRP)	Puerto Rico	KT966974	KT967012	KT967050	KT966936	KT966913
Thespesia lampas (Cav.) Dalzell [*Azanza lampas* (Cav.) Alef.]	TL1	*M. Balick and W. Nanakorn 3440* (MO)	Thailand	KT966966	KT967005	KT967042	KT966928	KT966906
Thespesia patellifera Borss.Waalk.	TP3	*J. R. Croft et al. LAE 68767* (L)	Papua New Guinea	KT966987	KT967025	KT967063	KT966949	KT966923
Thespesia populnea (L.) Sol. ex Corrêa	T24	*J. D. Ackerman et al. 4537* (UPRRP)	Hawaii	KT966982	KT967020	KT967058	KT966944	KT966919
	T94	*P. A. Gonzalez 1202-1* (HAJB, UPRRP)	Cuba	KT966983	KT967021	KT967059	KT966945	KT966920
	T98	*D. Santamaría 8868* (INB)	Costa Rica	KT966984	KT967022	KT967060	KT966946	KT966921
	T105	*F. Areces and V. Vega 668* (UPRRP)	USA	KT966985	KT967023	KT967061	KT966947	KT966922
Thespesia populneoides (Roxb.) Kostel.	TPo1	*A. Kanis 1987* (MO)	Australia	KT966986	KT967024	KT967062	KT966948	–

Table . *Continued*

Species	Code	Voucher	Origin	GenBank accession number				
				trnk/matK	*ndhF*	*rpl16*	ITS	*CesA1*
Thespesia thespesioides (Benth.) Fryxell [*Azanza thespesioides* (Benth.) F.Areces]	TTh			AY321161	U55326	AF384625	U56780	AF376044
	TTh1	*I. Cowie and J. Wendell 4233* (CANB)	Australia	KT966965	KT967004	KT967041	KT966927	KT966905
Urena lobata L.	UL	*F. Areces et al. 608* (UPRRP)	Puerto Rico	KT966995	KT967033	KT967071	KT966957	–

Appendix 2 Fossils of Malvaceae–Malvoideae used in dating analyses.

Fossil taxon	Structure	Location	Period/Epoch	Age (Myy)	Affinity	Reference	Age reference
Hibiscoxylon nyloticum Kräusel	Wood	Egypt	Late Cretaceous	88–66	Core Malvoideae	Kräusel (1939)	
Malvaciphyllum macondicus M. Carvahlo	Leaves	Colombia	Mid- to Late Palaeocene	60–58	Eumalvoid clade	Carvalho *et al.* (2011)	
Malvacipollis sp. 2	Pollen	USA	Late Palaeocene	58–56	Core Malvoideae	Gaponoff (1984)	
Malvacipollis tschudyi Frederiksen	Pollen	Cuba	Mid-Eocene	45	Core Malvoideae	Graham *et al.* (2000)	Graham (2003)
Malvacearumpollis sp.	Pollen	Australia	Late Eocene to Early Oligocene	37–30	Tribe Malveae	MacPhail & Truswell (1989)	
Malvacearumpollis mannanensis Wood	Pollen	Australia	Late Oligocene to Early Miocene	30–16	Tribe Malveae	Wood (1986)	
Malvocarpon clarum Hollick	Fruit	Puerto Rico	Mid-Oligocene	29	Tribe Malveae (*Abutilon* alliance)	Hollick (1928)	Graham (2003)
Abutilon sp.	Pollen	Puerto Rico	Mid-Oligocene	29	Tribe Malveae (*Abutilon* alliance)	Graham & Jarzen (1969)	Graham (2003)
Malvacipollis argentina Zamaloa & Romero	Pollen	Argentina	Late Oligocene to Mid-Miocene	28–14	Tribe Malveae	Zamaloa Mdel & Romero (1990)	Zamaloa Mdel (2000)
Baumannipollis chubutensis Barreda	Pollen	Argentina	Early Miocene	23–16	Tribe Malveae	Barreda (1993)	
Malvacearumpollis bakonyensis Nagy	Pollen	Hungary	Early Miocene	18–17	Tribe Malveae	Nagy (1962)	

Table . *Continued*

Fossil taxon	Structure	Location	Period/Epoch	Age (Myy)	Affinity	Reference	Age reference
Malvapantocolporites silvinites Mautino, Cuadrado & Anzótegui	Pollen	Argentina	Mid Miocene	16–11	Tribe Malveae	Mautino, Cuadrado & Anzótegui (2004)	
Echiperiporites estelae Germeraad, Hopping & Muller	Pollen	Venezuela	Mid- to Late Eocene	45–34	Tribe Hibisceae	Germeraad *et al.* (1968)	
Echiperiporites estelae	Pollen	Borneo	Early Miocene	20	Tribe Hibisceae	Germeraad *et al.* (1968)	
Hampea/Hibiscus (*Echiperiporites* sp.)	Pollen	Panama	Early Miocene	19.5–19	Tribe Hibisceae	Graham (1988)	Jaramillo *et al.* (2014)
Hampea/Hibiscus	Pollen	Panama	Mid- to Late Miocene	12–8.5	Tribe Hibisceae	Graham (1991)	Jaramillo *et al.* (2014)
Hampea/Hibiscus	Pollen	Mexico	Mid-Pliocene	4–3	Tribe Hibisceae	Graham (1976)	Graham (1994)
Gossypium tomentosum Nutt. ex Seem.	Leaves	Hawaii	Pleistocene	0.135–0.120	Tribe Gossypieae (*Gossypium* subgenus *Karpas*)	Woodcock & Manchester (1998)	