DO CHROMOSOME NUMBERS REFLECT PHYLOGENY?
NEW COUNTS FOR BOMBACOIDEAE AND A REVIEW OF MALVACEAE S.L.¹

RAFAELA C. MARINHO², CLESNAN MENDES-RODRIGUES³, FRANCISCO BALAO⁴, PEDRO L. ORTIZ⁴, JÚLIA YAMAGISHI-COSTA³, ANA M. BONETTI², AND PAULO E. OLIVEIRA³,⁵

¹Instituto de Genética e Bioquímica, Universidade Federal de Uberlândia, Uberlândia-MG 38400-902, Brazil; ²Instituto de Biologia, Universidade Federal de Uberlândia, Uberlândia-MG 38400-902, Brazil; and ³Departamento de Biología Vegetal y Ecología; Universidad de Sevilla; Apdo. 1095, E-41080 Sevilla, Spain

• Premise of the study: Whole genome duplication (WGD) and specific polyploidy events marked turning points for angiosperm genome structure and evolution. Therefore, cytogenetic studies of polyploidy-prone groups such as the tropical Malvaceae and plant formations such as as the Brazilian Cerrado have gained further importance. We present new chromosome counts for Cerrado Bombacoideae and revised chromosome numbers for the Malvaceae s.l., compare these between subfamilies, and relate them to phylogenetic signal.

• Methods: We studied the chromosome number of Eriotheca candolleana, E. gracilipes, E. pubescens, Pachira glabra, Pseudobombax longiflorum, and P. tomentosum. We also compared Eriotheca species ploidy levels using flow cytometry. We compiled chromosome numbers for 557 species of Malvaceae s.l., including 37 Bombacoideae species. We included this information in a phylogenetic reconstruction based on chloroplast matK-trnK DNA to evaluate chromosome evolution of the Malvaceae s.l. and the Bombacoideae in particular.

• Key results: The Cerrado Bombacoideae presented consistently high chromosome numbers. Numbers for Eriotheca species were among the highest and varied among populations. Flow cytometry analyses showed similar 1Cx DNA for all cytotypes and indicated neopolyploidy. Chromosome numbers differed between subfamilies, with the lowest numbers in the Malvoideae and Byttnerioideae and the highest in Tilioideae. Chromosome numbers had significant phylogenetic signal for Bombacoideae but not for Malvoideae or Malvaceae s.l.

• Conclusions: Clearly distinct chromosome numbers allied to monophyly provide some support for a circumscription of the Bombacoideae and distinction within the Malvaceae. The phylogenetic signal for chromosome number supports the idea of an ancient WGD and further neopolyploidy events as important evolutionary trends for the Bombacoideae.

Key words: Bombacoideae; Cerrado; chromosome number; Eriotheca; Malvaceae; phylogeny; polyploidy.

Cytogenetic data have been used in taxonomic studies since first half of last century. Such studies have provided insights into evolution, species segregation, and population features, although extensive studies for tropical plants are still lacking (Ehrendorfer, 1970; Guerra, 1990; Husband et al., 2013; Weiss-Schneeweiss and Schneeweiss, 2013). Chromosome counts and analyses are relatively quick, simple, and inexpensive tools that allow the study of karyotype diversity and the identification of polyploidy and aneuploidy events, which can explain species boundaries and contribute to taxonomic studies (Guerra, 2008). Improved techniques and phylogenetic analyses have somewhat revived interest in cytological data (Soltis et al., 2009), helped confirm the circumscription of taxa, and provided researchers with an understanding of the evolutionary process in many groups (Soltis and Soltis, 2000; Escudero et al., 2012). Identification of WGD events as turning points for the evolution of the angiosperms (Jaillon et al., 2007; AGP, 2013) has resulted in further interest in cytogenetic studies (Soltis et al., 2014).

Polyploidy is a frequent and important phenomenon in plant evolution (Soltis and Soltis, 2000; Soltis et al., 2009) and may lead to rapid adaptive change and speciation (Ramsey and Schemske, 1998). Different estimates indicate that many speciation events in the angiosperms are associated with polyploidy (Otto and Whitton, 2000). WGD seems to have been important for the evolution of the flowering plants, and even species with small genome size, such as Arabidopsis thaliana, show evidence of genome duplication (Soltis and Soltis, 2013). Different cytotypes may coexist in a population, resulting in ecological and evolutionary variation, although the polyploids themselves may diversify at lower rates (Mayrose et al., 2011). Hence, understanding these cytotypes and their distribution is important to explain the evolutionary process of different taxa (Forni-Martins et al., 1995; Balao et al., 2009; Escudero et al., 2012).

Since polyploidy is very common, some authors argue that most angiosperms are actually polyploid in origin (Masterson,
1994), with chromosome numbers of most species much higher than the basic number reckoned for the group. The ancestral basic number is usually estimated from a general profile of chromosome counts in a group of organisms, and from this ancestral number, it is possible to analyze the variation and evolution of ploidy levels (Soltis et al., 2014). Based on the distribution of chromosome counts and frequency of odd and even numbers, it has been estimated that the basic number for the angiosperms would be smaller than x = 13, much lower than most extant plants (Grant, 1963; Masterson, 1994; Soltis and Soltis, 2013).

Cytogenetic data are readily available for many temperate plants but neotropical taxa are still incompletely studied, so that the evolution of ploidy levels in many groups is still poorly understood (Guerra, 1990; Husband et al., 2013). Among neotropical environments, the Cerrado is the second largest biome and has been identified as a “hot spot” for conservation since less than 20% of the original area remains still pristine (Myers et al., 2000). It is the largest continuous neotropical savanna area, presenting a mosaic of plant formations with scattered trees on a grass matrix, under a markedly seasonal climate, with dry winters and summer rains (Ribeiro and Walter, 2008). Recent fires have driven the evolution of different angiosperm groups in the area (Simon et al., 2009). The Cerrado supports notable plant diversity (Ratter et al., 1997), which has resulted from in situ diversification and speciation, but to what extent polyploidy may have been involved remains an open question (Forni-Martins and Martins, 2000). Polyploidy has been reported for some Cerrado groups and seems to be associated with the extreme environmental conditions of this biome (Morawetz, 1986).

The family Bombacaceae includes tropical trees that are considered mostly paleopolyploid, with small and numerous chromosomes (Morawetz, 1986; Gibbs et al., 1988; Baum and Oginuma, 1994). Phylogenetic studies based on molecular data both from chloroplast and nuclear DNA, have resulted in the Bombacaceae being treated as a subfamily, the Bombacoideae, within the Malvaceae sensu latu (Baum et al., 1998; Alverson et al., 1999; Bayer et al., 1999; APG III, 2009). Recent molecular analyses (Baum et al., 2004; Duarte et al., 2011) have confirmed the monophyly of the Malvoideae and Bombacoideae and included these two subfamilies in the clade Malvatheca, clearly distinguished from the other families now included in the broader Malvaceae.

Polyploidy is common in the subfamily Bombacoideae. In the genus Adansonia, the baobab trees of Africa and Madagascar, up to $2n = 160$ chromosomes have been reported (Baum and Oginuma, 1994), although lower numbers also occur and species can be distinguished by their chromosome number (Pettigrew et al., 2012). As in other bombacoids, variable numbers have been reported for some Adansonia species (Miègue, 1974), but in some cases they may have been erroneous due to methodological problems, and more conservative and less variable numbers have been confirmed (Baum and Oginuma, 1994). In Brazil, even higher chromosome numbers have been described for Eriotheca trees. Eriotheca pubescens (Mart. & Zucc.) Schott & Endl. is polyploid with up to $2n = 276$ (Maglio et al., 1984; Oliveira et al., 1992; Forni-Martins et al., 1995; Mendes-Rodrigues et al., 2005), while Eriotheca gracilipes (K. Schum.) A. Robyns is mostly diploid but with a relatively high chromosome number ($2n = 92$; Maglio et al., 1984; Oliveira et al., 1992; Forni-Martins et al., 1995). However, some variation in chromosome numbers has been observed, and hitherto no other method has been used to assess genome size or provide further details on the chromosome evolution of these Eriotheca species.

 Possibly because of these difficulties, the high chromosome numbers and polyploid series in the Bombacoideae have not been taken into account in the phylogenetic reconstructions of the Malvaceae. The present study provides new chromosome counts for some Cerrado Bombacoideae, validated in some cases by flow cytometry, together with a review of the chromosome numbers for the Malvaceae sensu lato. We further used this information to evaluate the phylogenetic signal of chromosome information for the evolution of the Malvaceae as a whole and the Bombacoideae in particular, discussing a putative WGD as the source of innovation and diversification in these plants.

**MATERIALS AND METHODS**

**Cytogenetic analysis**—We studied chromosome number from root meristems of six Cerrado Bombacoideae species: Eriotheca candolleana (K. Schum.) A. Robyns, Eriotheca gracilipes (K. Schum.) A. Robyns, Eriotheca pubescens (Mart. & Zucc.) Schott et Endl., Pachira glabra Pasq., Pseudobombax longiflorum (Mart. & Zucc.) A. Robyns, and Pseudobombax tomentosum (Mart. & Zucc.) Robyns. We collected seeds from plants in different areas of Minas Gerais (MG) and Goiás (GO) states. Seeds of *P. glabra*, *E. candolleana*, *P. longiflorum*, and *P. tomentosum* were collected along the road between Uberlândia, Minas Gerais (MG) and Belo Horizonte, MG (BR452–BR262). Since *E. gracilipes* and *E. pubescens* vary in breeding system, giving rise to either polyploidy and apomictic, or monoapomictic and sexual individuals (see details in Mendes-Rodrigues et al., 2005, 2010, 2011; Marinho et al., 2014), chromosomes were counted for each individual sampled to identify possible cytotype differences in the species or population (see Table 1 for the number of individuals). Seeds from polyploidy individuals of *E. pubescens* were collected along the road from Uberlândia, MG to Brasília, Distrito Federal (BR050) and also in Caldas Novas, Goiás (GO), while monoapomictic individuals were collected in the region of Cristalina, GO. Polyembryonic individuals of *E. gracilipes* were collected in Caldas Novas, GO and considered monoapomictic ones were collected in Uberlândia, MG. Vouchers of the studied plants were deposited in the Herbarium Uberlândense (HUFU) of the Universidade Federal de Uberlândia (E. pubescens), monoapomictic individuals: Herbarium record HUFU 50731; *E. pubescens*, polyploidy individuals: HUFU 25854; *E. gracilipes*, monoapomictic individuals: HUFU 25856; *E. gracilipes*, polyploidy individuals: HUFU 50733; *E. candolleana*, HUFU 50730; *P. glabra*, HUFU52182; *P. longiflorum*, HUFU 56348.

The collected seeds were germinated on vermiculite in plastic boxes and watered as necessary. The radicles were pretreated either in a saturated solution of paradichlorobenzene (PDB) for 4 h at 16–18°C or for 24 h at 4°C, or for 0.002 M 8-hydroxyquinoline (8-hq) for 4 h at 16–18°C or for 24 h at 4°C. After the pretreatment, the radicles were fixed in Carnoy’s solution (3:1 ethanol–acetic acid v/v) for at least 24 h at room temperature and stored in 70% ethanol at −20°C. Root meristems were collected for cytogenetic analyses at 8-hq for 4 h or for 24 h at 4°C or for 0.002 M 8-hydroxyquinoline (8-hq) for 4 h at 16–18°C or for 24 h at 4°C.

**Table 1. Chromosome numbers for some Bombacoideae of the Cerrado region. Number of counted cells and ploidy level of individuals in each species.**

<table>
<thead>
<tr>
<th>Species</th>
<th>No. sampled individuals</th>
<th>No. progeny studied</th>
<th>No. counted cells</th>
<th>Ploidy 2n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eriotheca candolleana</td>
<td>2</td>
<td>2</td>
<td>20</td>
<td>2x = 92</td>
</tr>
<tr>
<td>Eriotheca gracilipes</td>
<td>3</td>
<td>5</td>
<td>20</td>
<td>2x = 92</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>11</td>
<td>25</td>
<td>6x = 276</td>
</tr>
<tr>
<td>Eriotheca pubescens</td>
<td>7</td>
<td>13</td>
<td>37</td>
<td>4x = 184</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>12</td>
<td>22</td>
<td>6x = 276</td>
</tr>
<tr>
<td>Pachira glabra</td>
<td>1</td>
<td>6</td>
<td>20</td>
<td>2x = 88</td>
</tr>
<tr>
<td>Pseudobombax longiflorum</td>
<td>3</td>
<td>8</td>
<td>22</td>
<td>2x = 88</td>
</tr>
<tr>
<td>Pseudobombax tomentosum</td>
<td>6</td>
<td>10</td>
<td>20</td>
<td>2x = 88</td>
</tr>
</tbody>
</table>
were rinsed, softened in HCl 5 N for 20 min and rinsed three times in distilled water. The meristems were then squashed in 45% acetic acid to spread the cells. The slides were frozen in liquid nitrogen to remove the coverslip, stained with 2% Giemsa (Guerra, 1983) and then sealed with Entellan mounting solution (Merck Millipore, Darmstadt, Germany). At least 20 metaphase cells were observed using an Olympus BX 51 microscope to count chromosomes for each species and cytotype. The best metaphase plates were photographed using an Olympus DP70 digital camera.

Flow cytometry—The absolute nuclear DNA content of 124 individuals (seedlings) was determined by flow cytometry. We used 25–37 seedlings from 5 to 7 polyembryonic and monoembryonic trees. For E. gruillieri (Vacu) from E. gracilipes (2n = 2x = 24) were collected from a polyembryonic population of Caldas Novas, GO and from a monoembryonic population in Uberlândia, MG. For E. pubescens, seeds from both polyembryonic and monoembryonic trees were collected in Cristalina, GO. Fresh leaf material from 1- to 2-mo-old seedlings grown in the greenhouse was used. For polyembryonic seeds that produced more than one seedling, one seedling per seed was selected at random for measurement.

Nuclear suspensions for flow cytometry measurements were prepared using a protocol adapted from Doležel et al. (1989). Approximately 25–50 mg of leaf tissue was chopped with a razor blade on a Petri dish (kept on ice) containing a protocol adapted from Doležel et al. (1989). Approximately 25–50 mg of leaf tissue was chopped with a razor blade on a Petri dish (kept on ice) containing 1 mL of LB01 buffer. The resulting suspension was kept on ice for 5–10 min and then a further 0.5 to 1.5 mL of the buffer was added to improve fluidity. The solution was filtered through a 30-mm mesh CellTric disposable filter (Partec GmbH, Münster, Germany), and mixed with 50 µL of propidium iodide (50 µg/mL) and 50 µL of RNase (50 µg/mL). The quality and the number of nuclei were dependent on solution fluidity. The leaves of Eriotheca produced a lot of gelatinous mucilage after chopping, which made it difficult to filter the material and apparently decreased the quality and number of nuclei obtained. For this reason, it was necessary to add more buffer before filtering. Until analysis, samples were kept on ice and protected from light. Flow cytometry measurements were taken using a Coulter CYTOMICS FC500-MPL (Beckman Coulter, Fullerton, California, USA) equipped with a 20-mW argon-ion laser at 488 nm. For each sample, we made 2–5 measurements on different days for all analyses.

We used Pisum sativum ‘Citriad’ with 9.09 pg of 2C nuclear DNA (Doležel et al., 1998) as the primary internal standard to determine DNA amount. For some samples, we also used Zea mays ‘CE 777’ with 5.43 pg of 2C nuclear DNA (Lysak and Doležel, 1998). In the hexaploid Eriotheca material, peaks overlapped with the Pisum standard, so we used known tetraploid specimens of Dianthus heteriori as a secondary standard (Baloo et al., 2009), which was recalibrated against P. sativum. The 2C genome size for this secondary standard was recalculated to DB 336/06 (n = 14; mean ± SE = 3.52 ± 0.02 pg; CV = 2.52%); to DB 236/06 3 (n = 13; mean ± SE = 3.60 ± 0.01 pg; CV = 2.38%); to DB 236/06 1 (n = 13; mean ± SE = 3.60 ± 0.01 pg; CV = 2.52%); and to DB263/07 3 (n = 13; mean ± SE = 3.17 ± 0.01 pg; CV = 1.72%). Peak means were established through manual gating using the software WinMDI 2.9 (Scripps Research Institute, La Jolla, California, USA; http://facs.scripps.edu/wm29w98.exe).

Overall, 399 flow cytometry measurements were performed. All measurements were used for analyses (i.e., within-plant extreme values were not dismissed; we found a variation of 2–11% from mean). Absolute DNA content was calculated for each ploidy level, tree, and seedling, and monoploid genome size (1Cx) was estimated as the amount of nuclear DNA divided by putative ploidy level. The 2C and 1Cx DNA and the coefficient of variation (scale response) were compared using generalized linear models (GLM) with Gaussian link function (Crawley, 2007).

Chromosome numbers in the Malvaceae s.l.—Chromosome numbers in the Malvaceae s.l. were compared using chromosome data available in literature and in the Index to Plant Chromosome Numbers (http://www.tropicos.org/Project/IPCN). We attempted to verify the original literature to avoid inconsistent data and excluded from the quantitative analyses any ambiguous numbers or those for which we could not check the original studies (only eight of 557 records, see Appendix S1 in Supplemental Data with the online version of this article). Mean chromosome numbers were obtained for each subfamily (sensu Baum et al., 2004; Duarte et al., 2011).

The mean chromosome numbers for each subfamily of the Malvaceae were analyzed with a nonphylogenetic generalized linear model (GLM) with Poisson loglinear distribution (Crawley, 2007) using the GLZM module in the program SPSS 17.0 (SPSS, Chicago, Illinois, USA) and a type III test. The means were compared pairwise with a least significant difference test. For the Malvaideae and Bombacoidaeae, the frequency distributions of chromosome numbers were analyzed in greater detail. Chromosome number classes were constructed using the Sturges rule (Sturges, 1926), and the differences between frequency distribution of classes of chromosome numbers were evaluated using a Kolmogorov–Smirnov (D) test (Sokal and Rohlf, 1981).

Phylogeny of Malvaceae and chromosome evolution—To understand the tempo and mode of chromosome evolution, we used phylogenetic comparative methods. Chloroplast matK-trnK DNA sequence data of a Malvaceae s.l. species and closely related taxa were used to construct a phylogeny. Accessions of Malvaceae s.l. were retrieved from GenBank (see online Appendix S2). We used 14 sequences of Cistaceae, Dipterocarpaceae, Muntingiaceae, and Thyme-laceae as outgroups. A data matrix for 201 accessions was aligned using the ClustalW algorithm implemented in the program Geneious 4.7.6 (Drummond et al., 2011). Phylogenetic analyses and divergence time estimates were performed simultaneously using the Bayesian uncorrelated log-normal method implemented in the program BEAST 1.6.1 (Drummond and Rambaut, 2007). Three independent BEAST runs were performed for 40 million generations each, using a birth–death prior for branching rates. Phylogenetic Markov chain Monte Carlo (MCMC) analyses were conducted under the GTR+ G model, which was selected based on Bayesian information criterion (BIC) values calculated in the program MrDol test2 (Darriba et al., 2012). We employed three independent calibration points. First, we calibrated the stem node of the Mal-voideae using the age of the earliest known fossil from the Late Paleocene. We applied a log-normal prior to this node, using a zero offset of 58 million years (Myr), which corresponds to the late Paleocene. Second, we calibrated the stem age of Cistaceae based on previous dating analyses (Guzmán and Vargas, 2009). We applied a normal age prior to these nodes with a mean of 14.20 Myr. Additionally, we applied a normal distribution with a mean of 23.95 Myr (as estimated by Guzmán and Vargas, 2009) to calibrate the node corresponding to the divergence of Cistaceae and its sister family Dipterocarpaceae. Adequacy of sampling and run convergence for the analyses was assessed using the effective sample size diagnostic in the program TRACER 1.5 (Rambaut and Drummond, 2007).

The phylogenetic signal in chromosome numbers (measured as logarithm of the maximum chromosome count) in a Malvaceae s.l. tree with 83 tips (only species with reliable known chromosome counts) was assessed using Pagel’s lambda (Pagel, 1999). Additionally, the ancestral states of chromosome number along the branches of a tree were estimated using a maximum likelihood approach (Revell, 2013) as implemented in the package phytools (Revell, 2012) in the R software ver. 3.0.1 (R Development Core Team, 2011). To investigate changes in chromosome number, we used the same approach as Escudero et al. (2012), based on Ornstein–Uhlenbeck models of trait evolution using the R packages matice (Hipp and Escudero, 2010) and ouch (King and Butler, 2009). Models for transitions in chromosome numbers occurring in (1) Malv- theca (Bombacoidaeae + Malvoideae), (2) Bombacoidaeae, and (3) Malvoideae were tested and compared by the corrected Akaike’s information criterion (AICc) and the Bayesian information criterion (BIC).

RESULTS

The Bombacoidaeae species studied here presented high chro- mosome numbers (Fig. 1, Table 1). Only the chromosome num- ber for Pseudobombax tomentosum and Eriotheca candolleana represented novelties at the species level. The other counts con- firmed previous studies, but we found new cytotypes for Erio- theca gracilipes (2n = 6x = 276) and Eriotheca pubescens (2n = 4x = 184) that had not been been previously recorded. Some of the metaphasic plates analyzed for species of the genera Erio- theca, Pachira, and Pseudobombax are presented (Fig. 1).

The flow cytometry analyses showed interesting results for the Eriotheca species and populations (Table 2). The 1Cx DNA was similar between species and cytotypes (Wald χ² = 0.270, df = 3, P = 0.996), which confirmed chromosome counts. Values of 2C DNA were simple multiples of this common value and supported the view that cytotypes are of neopolyploid origin. Although relatively few nuclei were measured per sample, variation was also low, and the results were consistent. There were
Fig. 1. Mitotic metaphasic plates from root meristem cells of some Malvaceae-Bombacoideae of the Brazilian Cerrado. (A) *Eriotheca candoleana* 2n = 2x = 92; (B) *Eriotheca gracilipes* 2n = 2x = 92 (Uberlândia-MG); (C) *Eriotheca gracilipes* 2n = 6x = 276 (Caldas Novas-GO); (D) *Eriotheca pubescens* 2n = 4x = 184 (Cristalina-GO); (E) *Eriotheca pubescens* 2n = 6x = 276 (Rodovia BR 050 and Caldas Novas-GO); (F) *Pachira glabra* 2n = 2x = 88; (G) *Pseudobombax longiflorum* 2n = 2x = 88; (H) *Pseudobombax tomentosum* 2n = 2x = 88. Scale bar = 10 µm.
The lowest chromosome numbers in Bombacoideae were found in Bombax insigne (2n = 18), endemic to the Indian Adaman Islands, and Pachira macrocarpa (2n = 26) from China. The highest numbers were recorded for the Eriotheca species of Brazilian Cerrado, with up to 276 chromosomes. There were clear differences in the means of chromosome numbers between some of the subfamilies of Malvaceae s.l. (Fig. 2). The lowest number was found for the Byttnerioideae and Malvoideae (2n = 10) and the highest for the Tilioideae (2n = 328). No chromosome count for the ninth subfamily, Brownlowioideae, was found in the surveyed literature. Despite some overlap, the Malvoideae and Bombacoideae, which have been included in the Malvotheca clade, presented clearly distinct frequencies of chromosome numbers (Kolmogorov–Smirnov test $D = 0.6021$, $P < 0.05$, Fig. 3). Similar analyses including the ambiguous or nonvalidated records (marked with asterisks in Appendix S1) gave similar results (not shown).

The phylogenetic analyses carried out in BEAST produced a phylogeny with moderate to high support (Fig. 4; online Appendix S3). The divergence time of the entire Malvaceae s.l. was 117.3 Myr (95% highest posterior density [HPD] = 91.2–154.2 Myr), which falls in the Lower Cretaceous. Bombacoideae appeared as a monophyletic group (including Ochroma) sister to the Malvoideae and apart from the other Malvaceae s.l. The crown node of Bombacoideae dates to the Paleocene-Late Eocene, around 60 Myr (HPD 42.7–75.7), but present-day genera diverged much later (less than 35 Myr). Based on this phylogenetic analysis, the phylogenetic signal for the chromosome number was strong ($\lambda = 0.77$, $P < 0.0001$) for Malvaceae s.l.,

### Table 2. Genome size estimates by flow cytometry for E. gracilipes and E. pubescens (Bombacoideae–Malvaceae) progeny from trees with different ploidy levels. Values are means ± SE. Means followed by different letters differed based on least significant difference ($P < 0.05$).

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Diploid</th>
<th>Hexaploid</th>
<th>Tetraploid</th>
<th>Hexaploid</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$ trees sampled</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>$N$ seedlings sampled</td>
<td>25</td>
<td>36</td>
<td>37</td>
<td>26</td>
</tr>
<tr>
<td>$N$ replicates</td>
<td>86</td>
<td>107</td>
<td>120</td>
<td>86</td>
</tr>
<tr>
<td>Nuclei per replication</td>
<td>$1160.21 \pm 107.5$</td>
<td>$995.79 \pm 60.8$</td>
<td>$1898.39 \pm 101.2$</td>
<td>$925.14 \pm 54.9$</td>
</tr>
<tr>
<td>2C DNA (pg)</td>
<td>$3.68 \pm 0.01$ a</td>
<td>$10.48 \pm 0.09$ a</td>
<td>$6.91 \pm 0.01$ c</td>
<td>$10.23 \pm 0.02$ b</td>
</tr>
<tr>
<td>1Cx DNA (pg)</td>
<td>$1.84 \pm 0.004$ a</td>
<td>$1.75 \pm 0.01$ a</td>
<td>$1.73 \pm 0.003$ a</td>
<td>$1.70 \pm 0.003$ a</td>
</tr>
<tr>
<td>Coefficient of variation (%)</td>
<td>$3.10 \pm 0.08$ b</td>
<td>$3.62 \pm 0.08$ a</td>
<td>$2.31 \pm 0.07$ c</td>
<td>$2.75 \pm 0.09$ c</td>
</tr>
</tbody>
</table>

**Notes:** GLM statistics for 2C DNA: Ploidy (Wald $\chi^2 = 2270.67$, df = 3, $P < 0.001$), Trees nested in ploidy ($\chi^2 = 12.75$, df = 12, $P = 0.387$), and seedlings nested in trees ($\chi^2 = 1.56$, df = 48, $P = 1.000$); for ICx DNA; ploidy ($\chi^2 = 0.270$, df = 3, $P = 0.996$), trees nested in ploidy ($\chi^2 = 0.259$, df = 18, $P = 1.000$); and for coefficient of variation; ploidy ($\chi^2 = 7.17$, df = 3, $P < 0.001$), trees nested in ploidy ($\chi^2 = 50.68$, df = 12, $P < 0.001$), and seedlings nested in trees ($\chi^2 = 25.94$, df = 48, $P = 0.996$).

Fig. 2. Mean chromosome number (and standard error) in each subfamily of Malvaceae s.l. The subfamilies means show statistical differences based on Wald $\chi^2$ test (7623.36, df = 7, $P < 0.001$); means followed by different letters indicate differences based on least significant differences. No chromosome count for the ninth subfamily, Brownlowioideae, was found in the surveyed literature.
indicating that phylogeny can help to predict the chromosome number of as yet unstudied taxa. The ancestral states mapped on the phylogeny showed chromosome transitions during Malvaceae evolution and a consistent trend of higher numbers among the Bombacoideae, while some possibly aneuploid lower numbers were found in the Malvoideae (Fig. 5). Furthermore, AICc and BIC weights for the three models of chromosome number transitions indicated that transitions were significant for the Bombacoideae clade (0.963 and 0.962, respectively), while they were not significant for the Malvatheca (0.424 and 0.270, respectively), or the Malvoideae (0.424 and 0.270, respectively).

![Graph showing frequency of species or cytotypes with different chromosome counts in the subfamilies Malvoideae and Bombacoideae (Malvaceae s.l.).](image1)

**Fig. 3.** Frequency of species or cytotypes with different chromosome counts in the subfamilies Malvoideae and Bombacoideae (Malvaceae s.l.).

![Backbone of a phylogenetic chronogram of the *matK-trnK* region inferred using BEAST. Time scale in millions of years ago (Ma). Three fossils and the estimated dates of the Malvoideae divergence and Cistaceae-Dipterocarpaceae divergence were used to calibrate the analyses (see online Appendix S3 for the complete phylogenetic chronogram.).](image2)

**Fig. 4.** Backbone of a phylogenetic chronogram of the *matK-trnK* region inferred using BEAST. Time scale in millions of years ago (Ma). Three fossils and the estimated dates of the Malvoideae divergence and Cistaceae-Dipterocarpaceae divergence were used to calibrate the analyses (see online Appendix S3 for the complete phylogenetic chronogram.).
double those observed in the subfamily Malvoideae, the sister group in the Malvatha clade. Actually, with the exception of a single species of Tilioideae, which has been recorded with a surprisingly high chromosome number, the Bombacoideae stands clearly above the mean chromosome number for the

**DISCUSSION**

The data presented here confirmed that the Bombacoideae is highly polyploid (Baum and Oginuma, 1994), and most plants in the subfamily showed chromosome numbers that are around double those observed in the subfamily Malvoideae, the sister group in the Malvatha clade. Actually, with the exception of a single species of Tilioideae, which has been recorded with a surprisingly high chromosome number, the Bombacoideae stands clearly above the mean chromosome number for the
family. These cytological data provide some support for the monophyly and family rank of the group upheld in a recent syn-
opsis (Cheek, 2007).

In our survey, among the Bombacoideae, *Bombax insigne* is the only nonpolyploid species, with $2n = 18$ (Sinha and Mazumdar, 1993) together with *Pachira macrocarpa* with $2n = 26$ chromosomes (Chen et al., 2003). These two numbers are unusual in an otherwise highly polyploid group. The number obtained for *Pachira macrocarpa* is similar to that recorded for most diploid species of *Gossypium* (Malvaceae) studied so far (Wendel and Cronn, 2003). Most of the Malvaceae present chromosome numbers lower than those found for the Bomba-
coideae, indicating that the origin of these latter taxa represents an important transition in the group as a whole. Phylogenetic analyses clearly support the importance of chromosome num-
ber as a trait/signal for the evolution of the Bombacoideae. As a cautionary remark, our phylogenetic analyses of chromosome evolution explicitly assume a continuous distribution (chromo-
some numbers were log-transformed to improve linearity of response). However, because polyploid events result in instant doubling of chromosomes, the assumptions of the OU model are not strictly met. The chromosome transition observed in Bombacoideae probably represented a WGD event, which could have happened over 60 Myr ago (Ma). This estimate would coincide with the major increase in polyploid lineages at the Cretaceous–Tertiary (KT) boundary (Soltis et al., 2014; Vanneste et al., 2014) and possibly soon after the Gondwanan split (Cheek, 2007).

The chromosome number $2n = 86$ predominates in many Bombacoideae genera. Most of the large tropical *Ceiba* trees, except for one report for *C. pentandra* ($2n = 88$; Gill et al., 1979), have this number. The *C. pentandra* variant number also appears in some other species of other genera, such as *Spirotheca rosea*, *Adansonia grandieri*, *Pachira glabra*, *Pachira aquat-
ica*, *Pseudobombax longiflorum*, and *Pseudobombax tomento-
sum*. Somewhat lower numbers in addition to the $2n = 86$ have also been reported in *Adansonia madagascariensis* and *Ceiba speciosa* with $2n = 80–84$. All these variations may represent aneuploidy or simply errors due to the problems posed by small and numerous chromosomes. Such analytical mistakes or even misidentifications were common in early cytological studies and have hindered cytotaxonomic interpretation. Unusual num-
bers with no clear link to a polyploid series, as observed for some species of *Adansonia*, such as *A. digitata*, *A. grandieri*, and *A. gregorii*, were considered to be analytical mistakes due to the small size of the chromosomes in the group by Baum and Oginuma (1994). However, the results presented here for the neotropical *Eriotheca*, backed by flow cytometry analyses, confirm that variant counts in the studied species are actual cyto-
typic variation and fit expected values for polyploidy series. New studies also support relatively recent diploid to polyploid divergence in *Adansonia* (Pettigrew et al., 2012).

We found different cytotypes in two *Eriotheca* species, and both can be interpreted as incomplete polyploid series. The ba-
sic number proposed for the genus *Eriotheca* is $x = 46$ chromo-
somes (Forni-Martins et al., 1995), so that it is possible to view the cytotypes of *E. pubescens* as $2n = 4x = 184$ and $2n = 6x = 276$ chromosomes. As for *E. gracilipes*, previous recorded numbers were either $2n = 2x = 92$ or 96 (Oliveira et al., 1992; Forni-
Martins et al., 1995). A report of $2n = 210$ for *E. gracilipes* (Morawetz, 1986, not included in IPCN) was interpreted as a misidentification of a cytotype of *E. pubescens* (Oliveira et al., 1992). However, the present survey indicates that there may have been no misidentification, since individuals of *E. gracilipes* studied here showed both $2n = 2x = 92$ and $2n = 6x = 276$ chromosomes. Flow cytometry results supported this interpre-
tation since basic 1Cx DNA was similar between all *Eriotheca* species and cytotypes and 2C DNA values are simple multiples of the monoploid genome.

The cytotypes of these *Eriotheca* species showed different frequencies in the field. According to our sampling, the putative hexaploid cytotype in *E. pubescens* is widely distributed, while the tetraploid cytotype $2n = 4x = 184$ chromosomes is restricted to a single population in Cristalina-GO. The opposite occurs in *E. gracilipes*, for which hexaploid individuals were found only in a relatively limited area in Caldas Novas, GO (Mendes-Rodrigues, 2010).

Higher ploidy levels, such as the hexaploid cytotypes in *E. gracilipes* and *E. pubescens*, may be allopolyploids derived via hybridization between diploid and tetraploid sexual individuals followed by polyploidization. Alternatively, they may be auto-
polyploids that have arisen from reduced and unreduced gam-
etes (Ramsey and Schemske, 1998). Morphological similarity between the hexaploid ampollics and lower ploidy sexuals of each species to a certain extent support the latter pathway. En-
vironmental factors such as high temperature, herbivory, and low nutrient availability may influence unreduced pollen produ-
duction (Ramsey and Schemske, 1998), and since such factors are common Cerrado features, they could explain the origin of these highly polyploid individuals. However, these hexaploids are also apomicts and apomixis has been associated to with allopolyploids and asynchronies during embryogenesis due to hybrid origin (Carman, 1997).

The phylogenetic analyses indicate that these *Eriotheca* polypody events seem to be relatively recent (i.e., neopoly-
ploidy), occurring probably less than 10 Ma. Recent polyploidiza-
tion events involve cytotype formation and demographic establishment (Ramsey and Schemske, 2002), and populations of neopolyploids usually have mutations or special gene combi-
tations that are fixed even when phenotypic changes are small, due to reproductive isolation (Soltis et al., 2014). Such geno-
types may confer advantages to these individuals. In *Adanso-
nia*, the recently found diploid *A. kilima* has a restricted distribution while *A. digitata*, which seems to comprise recent neopolyploids, are much more widespread (Pettigrew et al., 2012).

Apolixis and polyembryony, often associated with poly-
ploidy (Carman, 1997; Koltunow and Grossniklaus, 2003), may also be advantageous traits since they allow greater re-
productive assurance (e.g., Santos et al., 2012) and may confer adaptive ability to habitats different than those occupied by dip-
loids (Ramsey and Schemske, 1998; Otto and Whittington, 2000). Such highly polyploid cytotypes associated with apomixis and polyembryony may have favored the distribution and adaptabil-
ity of *E. pubescens* (Oliveira et al., 1992; Mendes-Rodrigues et al., 2005; Mendes-Rodrigues, 2010). It is noteworthy that the association between apomixis and neopolyploidy in *Eriotheca* contrasts with the trend for apomixis to be associated with pa-
leopolyploidy (Ramsey and Schemske, 2002).

Chromosome numbers may also help us to interpret phyloge-
netic relationships. Recent molecular analyses of the Bomba-
coideae (Duarte et al., 2011) confirm its monophyly and indicate a division in three clades: the first with the genera *Bermouilla*, *Gyranthera*, and *Huberodendron*; the second with *Adansonia*, *Catostemma*, *Cavanillesia*, and *Scleronea*; and the third with *Bombax*, *Ceiba*, *Neobuchia*, *Pachira-Eriotheca*, *Pseudobombax*,
Rhodognaphalon, and Spirotheca. The latter clade indicates an affinity between Eriotheca and Pachira (Duarte et al., 2011). But there were some differences in chromosome number between species, and since the variable cytotypes in Eriotheca are novelties, duplication events may have an impact on the nuclear ITS genes used in the phylogenetic reconstruction by Duarte et al. (2011). Since plastid genes upheld the monophyly of Eriotheca, it may be necessary to reassess which plants/populations were used in the reconstruction and what may be the implications of the polyploid series described here to the phylogeny of the group.

The inclusion of both Ochroma and Patinoa in the Bombacoideae is also somewhat contentious. Based on the molecular analyses, these genera fall between the Bombacoideae and the Malvoideae (Duarte et al., 2011). We could not find any information for Patinoa, but reports for chromosome numbers for Ochroma pyramidalare $2n = 2x = 78$, 88, and 90. Although variable, these numbers are similar to those described for Bombacoideae and much higher than those described for the Malvoideae. On the basis of the phylogenetic relationship, we predict much lower numbers of the phylogenetic relationship, we predict much lower numbers of the Bombacoideae. The numbers

On the basis of the phylogeny and published chromosome counts, we can define two basic chromosome numbers for the Bombacoideae. The numbers $x = 44$ and 46 were found commonly among the species and explain most ploidy variation observed for the group. The variability in numbers that occur in some taxa are possibly due to aneuploidy, with loss or gain of one or a few chromosomes in the basic genome, resulting in $2n = 72, 86, 96,$ and 160. Strikingly, the chromosome numbers compiled here are much higher than those found in species in the Malvaceae and also most of the Malvaceae s.l. Chromosome number and possibly genome size have a clear phylogenetic signal and support the Bombacoideae as a monophyletic group. This group is possibly a result of an ancient whole genome duplication event that occurred around the KT boundary and of subsequent recent and independent polyploidization events. This polyploidy-prone evolution pattern of the Bombacoideae differs from that described for other species-rich groups (Escudero et al., 2012), in which polyploidy seems to have had a less important role.

LITERATURE CITED


