



OPEN ACCESS

EDITED BY

Daniel Pinero,
National Autonomous University of Mexico,
Mexico

REVIEWED BY

He Jian,
Beijing Forestry University, China
Chenyang Cai,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE

André
Marques
✉ amarques@mpipz.mpg.de
Gustavo Souza
✉ luiz.rodriguesouza@ufpe.br

RECEIVED 02 October 2025

REVISED 23 December 2025

ACCEPTED 31 December 2025

PUBLISHED 23 January 2026

CITATION

Luna G, Costa L, Pezzini FF, Karimi N,
Strijk JS, Carvalho-Sobrinho J, Colli-Silva M,
Marques A and Souza G (2026) How much
can reticulate evolution entangle plant
systematics? Revisiting subfamilial
classification of the Malvaceae clade
(Malvaceae) on the basis of phylogenomics.
Front. Plant Sci. 16:1717745.
doi: 10.3389/fpls.2025.1717745

COPYRIGHT

© 2026 Luna, Costa, Pezzini, Karimi, Strijk,
Carvalho-Sobrinho, Colli-Silva, Marques and
Souza. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

How much can reticulate evolution entangle plant systematics? Revisiting subfamilial classification of the Malvaceae clade (Malvaceae) on the basis of phylogenomics

Gustavo Luna¹, Lucas Costa^{1,2}, Flávia Fonseca Pezzini³,
Nisa Karimi⁴, Joeri Sergej Strijk⁵,
Jefferson Carvalho-Sobrinho^{6,7}, Matheus Colli-Silva^{8,9}, André
Marques^{10*} and Gustavo Souza^{1*}

¹Laboratory of Plant Cytogenetics and Evolution, Department of Botany, Federal University of Pernambuco, Recife, Pernambuco, Brazil, ²Laboratory of Evolutionary Ecology and Plant Genomics, Institute of Biology, State University of Campinas, Campinas, Brazil, ³Royal Botanic Garden Edinburgh, Edinburgh, United Kingdom, ⁴Department of Botany, University of Wisconsin – Madison, Madison, WI, United States, ⁵Alliance for Conservation Tree Genomics, Kuala Lumpur, Malaysia, ⁶Universidade Federal do Vale do São Francisco, Petrolina, Brazil, ⁷Universidade Federal Rural de Pernambuco, Recife, Brazil, ⁸Laboratory of Plants Morphotaxonomy and Evolution, Department of Botany, Federal University of Pernambuco, Recife, Pernambuco, Brazil, ⁹Royal Botanic Gardens, Richmond, United Kingdom, ¹⁰Department of Chromosome Biology, Max Planck Institute for Plant Breeding Research, Cologne, Germany

Reticulate evolution (RE), involving hybridization and related processes, generates network-like rather than strictly bifurcating relationships among lineages and can obscure phylogenetic relationships. Detecting ancient hybridization is particularly challenging, as genomic signals may erode over time. The Malvaceae clade (Malvaceae), marked by multiple paleopolyploidy events since its estimated origin 66 my, offers a useful model for examining RE. Its three subfamilies—Bombacoideae (with high chromosome numbers, mostly trees), Malvoideae (lower chromosome numbers, mostly herbs), and the recently described Matisoideae—show unresolved relationships, with several taxa of uncertain placement. We conducted a phylogenomic analysis of 69 Malvaceae species via complete plastomes, 35S rDNA cistrons, nuclear low copy genes and comparative repeatome data. Most of the datasets consistently resolved four clades: (I) Bombacoideae, (II) Malvoideae, (III) Matisoideae, and (IV) a heterogeneous assemblage including representatives of Malvoideae, Matisoideae and several *incertae sedis* taxa. Chromosome numbers were negatively correlated with repeatome diversity: Bombacoideae presented higher counts but lower repeat diversity, possibly reflecting slower repeat evolution associated with woody growth forms. In contrast, clades III and IV showed marked heterogeneity in both chromosome number and repeat

composition, which is consistent with a reticulate origin. Overall, our results show evidence of ancient hybridization and polyploidy in shaping *Malvatheca* evolution. These results highlight that reticulation and genome dynamics, rather than taxonomic boundaries alone, are central to understanding the diversification of *Malvatheca*.

KEYWORDS

bombacoideae, malvoideae, matisioideae, phylogenomics, polyploidy, reticulate evolution

Introduction

Reticulate evolution (RE) refers to processes in which genetic material is exchanged between lineages, producing network-like rather than strictly bifurcating phylogenetic relationships (Arnold, 1997; Gontier, 2015). It can arise through hybridization, horizontal gene transfer, or endosymbiosis, and challenges the traditional tree-like view of evolution (Mallet, 2007). Reticulation also complicates classification by blurring taxonomic boundaries and generating conflicting phylogenetic signals. In plants, where hybridization is common, genetic and phenotypic distinctions may be obscured, leading to uncertain or misleading taxonomic boundaries (Mallet, 2007).

Polyploidy (whole genome duplication, WGD), the presence of multiple complete genomes in a species, is closely linked to reticulation and is a recurrent feature of angiosperm evolution (Bowers and Paterson, 2021). Allopolyploidization, i.e., hybridization followed by genome doubling, has been recognized as a major driver of plant diversification (Paterson et al., 2004; Jiao et al., 2011; Vekemans et al., 2012). Notably, a burst of polyploidization events in flowering plants occurred near the Cretaceous–Paleogene (K–Pg) boundary (~66 Mya), coinciding with a mass extinction event (Vanneste et al., 2014; Fawcett and Van de Peer, 2010). The clustering of independent WGDs across angiosperms during this period suggests that polyploidy promoted survival and diversification by providing raw material for genetic novelty (Fawcett et al., 2009; Wood et al., 2009).

Studying RE in ancient lineages remains difficult because signals of hybridization and polyploidy are often eroded over time (Zhang et al., 2025). Additionally, disentangling processes such as incomplete lineage sorting, horizontal transfer, or genetic erosion is challenging (Mallet, 2007), and modeling ancient reticulation requires large datasets and sophisticated analyses. High-throughput sequencing has made these investigations feasible (Dodsworth et al., 2019). Target enrichment methods such as Hyb-Seq enable the analysis of thousands of low-copy nuclear loci (Karbstein et al., 2022), whereas genome skimming can recover plastid genomes (tracing maternal inheritance), ribosomal DNA (biparental nuclear

signal), and repetitive elements such as satellites and transposable elements (Dodsworth, 2015; Cavallini et al., 2019). Repeat abundance and sequence similarity offer complementary, alignment-free tools for phylogenetic reconstruction (Dodsworth et al., 2019; Vitales et al., 2020).

The *Malvatheca* clade (Malvaceae), which originated near the K–Pg boundary (~66 Mya; Hernández-Gutiérrez and Magallon, 2019; Cvetković et al., 2021), serves as a system for investigating these processes. It comprises Bombacoideae (17 genera, ~160 species), Malvoideae (78 genera, ~1,670 species), and the recently recognized Matisioideae (three genera, ~138 species) (Baum et al., 2004; Carvalho-Sobrinho et al., 2016; Cvetković et al., 2021; Colli-Silva et al., 2025). Matisioideae, formerly the tribe Matisieae, was elevated to subfamily rank by Colli-Silva et al. (2025) after consistent recovery of its monophyly, despite its morphological distinctiveness. While Matisioideae is a well-supported lineage, its position—sister to Bombacoideae or to a recircumscribed Malvoideae—remains poorly resolved, even with phylogenomic data. Malvoideae includes widely cultivated herbs such as cotton, hibiscus, and okra (Baum et al., 2004), whereas Bombacoideae consists mainly of tropical trees, including ecologically and culturally significant species such as the baobab (*Adansonia* L. sp.) and the kapok tree (*Ceiba pentandra* (L.) Gaertn.) (Wickens, 2008; Zidar and Elisens, 2009). Several genera, such as *Ochroma* Sw. and *Chiranthodendron* Larreat., placement remain uncertain due to conflicting morphological and molecular evidence. This phylogenetic uncertainty, coupled with the clade's history of polyploidy, makes *Malvatheca* a valuable system for studying reticulate evolution in ancient lineages.

The subfamilies within *Malvatheca* appear to have followed contrasting genomic trajectories shaped by postpolyploid diploidization. Bombacoideae is predominantly composed of species with high chromosome numbers ($2n = 86\text{--}276$; Costa et al., 2017), whereas Malvoideae ranges from $2n = 10\text{--}130$, with a modal value near $2n = 16$ (Tate et al., 2005). Genomic studies have reported signals of reticulate allopolyploidization in *Malvatheca* and his consequences to their phylogentic relationships (Hernández-Gutiérrez et al., 2022; Sun et al., 2024; Yang et al., 2025; Zhang et al.,

2025), but its consequences for genome architecture, or diversification still poorly known. To address this question, we analyze ribosomal DNA, plastome and repeatome data of 22 Bombacoideae, 34 Malvoideae, 7 Matisioideae, and 6 *incertae sedis* taxa under a comparative phylogenomic framework. Specifically, we ask the following questions: (i) Is there evidence of reticulate evolution in the group, and is it associated with taxa of uncertain placement? (ii) Are there repeatome signatures linked to ancient WGDs? (iii) Does lifeform correlate with the distinct diploidization pathways observed in Bombacoideae and Malvoideae?

Materials and methods

Data acquisition and repeatome characterization

We analyzed 69 species of the Malvatheca clade, including 34 Malvoideae, 22 Bombacoideae, 7 Matisioideae, and 6 *incertae sedis*. Two *Theobroma* species (Bytnerioideae) were included as outgroups. Species names and NCBI accession codes (when available) are listed in Table 1. Target-capture sequencing was used to expand the representation of Bombacoideae, following Costa et al. (2021). Not

TABLE 1 Species and accessions used.

Bombacoideae	<i>n</i>	1C (PG)	Access	Approach
<i>Adansonia digitata</i> L.	84	1.7	SRX19282309	WGS
<i>Adansonia grandidieri</i> Baill.	44	0.76	SRX24110595	WGS
<i>Adansonia gregorii</i> F.Muell.	44	0.68	SRX24110598	WGS
<i>Adansonia madagascariensis</i> Baill.	44	0.69	SRX24110597	WGS
<i>Adansonia perrieri</i> Capuron	42	0.71	SRX7064913	HybSeq
<i>Adansonia rubrostipa</i> Jum. & H.Perrier	44	0.65	SRX24110596	WGS
<i>Adansonia suarezensis</i> H.Perrier	44	0.74	SRX7064911	HybSeq
<i>Adansonia za</i> Baill.	44	0.63	SRX7064903	HybSeq
<i>Bombax ceiba</i> L. 1	46	1.62	SRX20143449	WGS
<i>Bombax ceiba</i> L. 2	46	1.62	SRX7064904	HybSeq
<i>Ceiba aesculifolia</i> (Kunth) Britten & Baker f.	–	–	PRJNA579976	HybSeq
<i>Ceiba pentandra</i> (L.) Gaertn. 1	43	1.75	SRX20143754	WGS
<i>Ceiba pentandra</i> (L.) Gaertn. 2	43	1.75	Pezzini et al., 2021	HybSeq
<i>Ceiba speciosa</i> (A.St.-Hil., A.Juss. & Cambess.) Ravenna	43	1.25	SRX18461125	WGS
<i>Eriotheca macrophylla</i> (K.Schum.) A.Robyns	138	4.77	SRX14481707	HybSeq
<i>Gyranthera amphibiolepis</i> W.Palacios	–	–	SRX14481737	HybSeq
<i>Huberodendron swietenoides</i> (Gleason) Duce	–	–	SRX14481725	HybSeq
<i>Neobuchia paulinae</i> Urb.	–	–	SRX14481682	HybSeq
<i>Pachira aquatica</i> Aubl	46	2.3	PRJNA1288544	HybSeq
<i>Pentaplaris davidsmithii</i> Dorr & C.Bayer	–	–	SRX14481702	HybSeq
<i>Pseudobombax croizatii</i> A.Robyns	–	–	SRX7064905	HybSeq
<i>Rhodognaphalon schumannianum</i> A.Robyns (<i>Rhodognaphalon mossambicense</i> (A.Robyns) A.Robyns)	–	–	PRJNA579976	HybSeq
<i>Scleronema micranthum</i> (Ducke) Ducke	–	–	SRX7064906	HybSeq
<i>Spirotheca rosea</i> (Seem.) P.E.Gibbs & W.S.Alverson	44	–	SRX14481726	HybSeq
Malvoideae				
<i>Abelmoschus esculentus</i> (L.) Moench	65	1.23	SRX21378022	WGS
<i>Abutilon amplum</i> Benth.	–	–	SRX5462903	WGS
<i>Abutilon fruticosum</i> Guill. & Perr.	21	–	ERX12138327	HybSeq

(Continued)

TABLE 1 Continued

Bombacoideae	<i>n</i>	1C (PG)	Access	Approach
Malvoideae				
<i>Abutilon theophrasti</i> Medik.	21	1.4	SRX9130922	WGS
<i>Alcea rosea</i> L.	21	–	ERX9653860	WGS
<i>Althaea officinalis</i> L.	21	–	SRX2855182	WGS
<i>Decaschistia byrnesii</i> Fryxell	–	–	ERX7170161	HybSeq
<i>Gossypium arboreum</i> L.	13	1.7	SRX5759203	WGS
<i>Gossypium herbaceum</i> L.	13	1.7	SRX9217821	WGS
<i>Gossypium raimondii</i> Ulbr.	13	0.9	SRX341462	WGS
<i>Hampea nutricia</i> Fryxell	13	–	SRX14481653	HybSeq
<i>Hibiscus brachysiphonius</i> F.Muell.	–	–	SRX5462993	WGS
<i>Hibiscus</i> × <i>rosa-sinensis</i> L.	42	–	ERX9653895	WGS
<i>Hibiscus syriacus</i> L.	40	–	SRX529352	WGS
<i>Hibiscus trionum</i> L.	28	–	SRR17686673	WGS
<i>Howittia trilocularis</i> F.Muell.	–	–	ERX7170167	Hybseq
<i>Kokia cookei</i> O.Deg.	12	0.57	SRX24024016	WGS
<i>Kokia drynarioides</i> (Seem.) Lewton	12	0.6	SRX3304973	WGS
<i>Kokia kauaiensis</i> (Rock) O.Deg. & Duvel	12	0.57	SRX24024021	WGS
<i>Kosteletzkya pentacarpos</i> (L.) Ledeb.	17	–	SRX12399957	WGS
<i>Lawrenzia densiflora</i> (Baker f.) Melville	–	–	SRX5463054	WGS
<i>Lavatera assurgentiflora</i> subsp. <i>glabra</i> (<i>Malva assurgentiflora</i> subsp. <i>glabra</i> (Kellogg) M.F.Ray)	–	–	SRX23987976	HybSeq
<i>Malva sylvestris</i> L.	21	1.5	ERX3149988	WGS
<i>Malva moschata</i> L.	21	1.4	ERX3149982	WGS
<i>Malva pusilla</i> Sm.	21	–	ERX5310142	WGS
<i>Napaea dioica</i> L.	14	8.5	SRX30129087	WGS
<i>Pavonia schiedeana</i> Steud.	–	–	SRX14481669	HybSeq
<i>Pavonia hastata</i> Cav.	28	–	ERX7170168	HybSeq
<i>Pavonia triloba</i> Guill. & Perr.	–	–	ERR5165833	HybSeq
<i>Pavonia urens</i> Cav.	–	–	ERR7622263	HybSeq
<i>Sida cardiophylla</i> F.Muell.	–	–	SRX5462742	WGS
<i>Sida clementii</i> Domin	–	–	SRX5462743	WGS
<i>Sida spinosa</i> L.	7	1	SRX5462745	WGS
<i>Thespesia populneoides</i> (Roxb.) Kostel.	13	–	SRX706235	WGS
Matisioideae				
<i>Matisia alchornifolia</i> Triana & Planch.	–	–	ERR7622266	HybSeq
<i>Matisia ochrocalyx</i> K.Schum.	–	–	ERR7622272	HybSeq
<i>Phragmothea ecuadorensis</i> W.S.Alverson	–	–	ERR7622267	HybSeq
<i>Phragmothea mammosa</i> W.S.Alverson	–	–	SRX14481713	HybSeq
<i>Quararibea asterolepis</i> Pittier	–	–	SRR15016217	HybSeq

(Continued)

TABLE 1 Continued

Bombacoideae	<i>n</i>	1C (PG)	Access	Approach
Matisioideae				
<i>Quararibea duckei</i> Huber	–	–	ERR14030174	HybSeq
<i>Quararibea funebris</i> (La Llave) Vischer	–	–	SRX14481652	HybSeq
Incertae sedis				
<i>Lagunaria patersonia</i> (Andrews) G.Don	–	–	SRX14481717	HybSeq
<i>Camptostemon schultzei</i> Mast.	–	–	SRX14481714	HybSeq
<i>Alyogyne huegelii</i> (Endl.) Fryxell	32	–	SRX14481721	HybSeq
<i>Chiranthodendron pentadactylon</i> Larreat.	–	–	SRX14481650	HybSeq
<i>Fremontodendron mexicanum</i> Davidson	–	–	SRX25649159	HybSeq
<i>Ochroma pyramidale</i> (Cav. ex Lam.) Urb.	42	2.2	SRX14481665	HybSeq
Outgroup (Byttnerioideae)				
<i>Theobroma cacao</i> L.	20	0.4	SRX23903616	WGS
<i>Theobroma grandiflorum</i> (Willd. Ex Spreng.) K.Schum.	20	0.51	SRX17564460	WGS

all species had reads with sufficient reads quality/amount to proceed with all analysis. Thus it was not possible to perform repeatome characterization based on most of the Matisioideae species (except *Pragmotheca mammosa* and *Quararibea funebris*).

Repeatome analyses were performed with the RepeatExplorer2 pipeline implemented in the Galaxy server (<https://repeateexplorer-elixer.cerit-sc.cz>; Novák et al., 2020), which clusters reads on the basis of sequence similarity. Two approaches were applied: (i) individual analyses, where each species was analyzed separately to characterize its repetitive fraction, and (ii) comparative analysis, where concatenated reads from all species were analyzed jointly to compare repeat composition across lineages.

For individual analyses, ~0.1× genome coverage per species was used (Table 1). For species with unknown genome sizes, values were estimated on the basis of closely related taxa. For the comparative analysis, reads were normalized to ~0.05× per species, yielding a total of 61,288,798 reads. In both cases, clustering was performed using 90% sequence similarity and 55% minimum overlap. The proportions of repeat lineages were calculated as the number of reads per cluster relative to the total number of reads analyzed, excluding chloroplast and mitochondrial sequences identified as potential contaminants.

Correlations between karyotypic, genomic and ecological traits

Following Schley et al. (2022), we treated the genome as a “community” and repetitive element lineages as “species,” allowing the application of community ecology metrics to genome composition. Repeat diversity was quantified via the Shannon diversity index (*H*; Shannon, 1948), which measures the probability that two randomly chosen repeat copies belong to the same lineage. Higher *H* values indicate greater repeat diversity.

For each species, diversity indices were calculated from repeat abundances obtained from the RepeatExplorer individual analyses. To test whether repeat diversity was associated with chromosome number, we performed a Pearson correlation between log₁₀-transformed Shannon index values and log chromosome counts across Malvatheca species via the stats package in R (R Core Team, 2019).

To further examine potential ecological correlates, we compiled chromosome numbers and growth habits (herbaceous vs. woody) for 84 species from the literature. Relationships between habit and chromosome number were visualized with boxplots constructed in PAST 4 (Hammer et al., 2001).

Plastome, rDNA, low copy nuclear genes and repeat-based phylogenies

Plastomes and ribosomal DNA (rDNA) sequences were assembled for all sampled species using a reference-based mapping approach in Geneious v6.0.3 (Kearse et al., 2012). *Theobroma cacao* (NC_014676.2) served as the reference for plastome assembly, while sequence JQ228369.1 was used for rDNA. Reads were mapped to their respective references using the “Map to reference” function, and consensus plastome sequences were aligned with MAFFT (Katoh and Standley, 2013).

The raw reads from each of the 69 species were mapped against 44,343 gene sequences from the *Ceiba pentandra* genome assembly (Shao et al., 2024) using the “Map to reference” function. These mapped genes were filtered based on mapping quality, phylogenetic resolution, and a coverage threshold requiring the gene to be present in at least 67 species. This filtering process yielded four genes suitable for phylogenetic analysis (CpUnG0025.1, CpUnG0488.1, CpUnG0595.1, CpUnG1740.1). A maximum likelihood (ML) phylogenetic tree was inferred from these four

genes, the plastome and the rDNA alignments using IQ-TREE software (v3.0.1). The resulting tree was visualized in FigTree (<https://tree.bio.ed.ac.uk/software/figtree/>). Topological incongruences between the plastid, rDNA, and nuclear gene trees were visualized using a circular tanglegram generated with the circize package (Gu et al., 2014) in R. Phylogenetic inference based on repeat abundances was performed following Dodsworth et al. (2015). The abundances of repeat classes obtained from comparative repeatome analyses were treated as quantitative characteristics. Parsimony analyses were conducted in PAST 4 (Hammer et al., 2001).

Reticulate evolution

Reticulate evolution was assessed by constructing phylogenetic networks in PhyloNet 3.8.2 (Wen et al., 2018). The input for this analysis consisted of the individual tree topologies from the complete plastome, rDNA, and four nuclear loci. We employed a two-step approach: (1) estimation of an underlying ML species tree from the set of input gene trees using the DeepCoalCount_tree algorithm (H_0); and (2) inference of a phylogenetic network by iteratively adding reticulations to this species tree with the DeepCoalCount_network algorithm to estimate the number of hybridization events (H_1 , H_2 , H_3 etc.). The inferred network was visualized using SplitsTree v6.4.17 (Huson and Bryant, 2006). To independently validate the crosslinking events proposed by network inference, we performed Patterson's D-statistic test (ABBA-BABA test) based on nucleotide site patterns (Green et al., 2010; Durand et al., 2011). For this, we used information from all loci (plastomes + rDNA + low-copy nuclear genes) analyzed here for representatives of the Malvatheca clade. We calculated the D value for specific taxonomic quartets ((P1, P2), P3, O) designed to test each gene flow hypothesis, where a significant deviation from zero ($D > 0$) indicates an excess of shared derived alleles between P2 and P3, consistent with introgression.

Results

Comparison between plastidial, rDNA and low-copy nuclear genes topologies

Clade I (Bombacoideae) was well-supported across all three topologies, with one key exception: *Huberodendron swietenoides* + *Gyranthra amphibiolepis* were absent from the rDNA topology, appearing instead within Clade III (Figure 1). Furthermore, on the plastid topology, the *incertae sedis* species *Fremontodendron mexicanum* was resolved as the earliest-diverging lineage of this clade (Figure 1).

Clade II, also well-supported, primarily comprises Malvoideae species. Exceptions include *Pavonia urens*, which was placed in Clade III (Matisioideae) on the plastid topology, and *Pavonia schiedeana*, which appeared in Clade IV in the low-copy nuclear and rDNA topologies (Figure 1). This clade also includes the *incertae sedis* species *Camptostemon schultzei*, *Howittia*

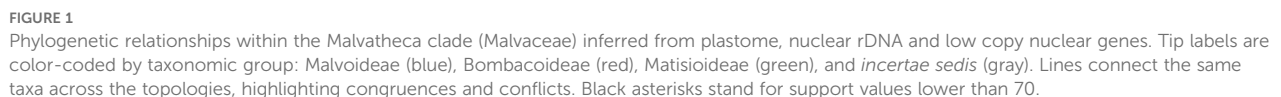
trilocularis, and *Pentaplaris davidsmithii*; the latter was placed here in the plastid topology, while *Fremontodendron mexicanum* was placed here in the rDNA topology (Figure 1).

Clade III (Matisioideae) was consistently well-supported (Figure 1). However, several species from other groups appeared within it depending on the topology: the *incertae sedis* species *Pentaplaris davidsmithii* and the Bombacoideae species *H. swietenoides* and *G. amphibiolepis* on the rDNA topology; the Malvoideae species *Pavonia urens* on the plastid topology (Figure 1). On the other hand, the Matisioideae species *Quararibea funebris* was placed in another clade (IV) across all analyses (Figure 1). Remarkable, the backbone of the subfamilies was variable depending on the dataset, with clade III (Matisioideae) being sister to Bombacoideae in rDNA topology or sister to Malvoideae in plastidial and low-copy nuclear gene topologies (Figure 1).

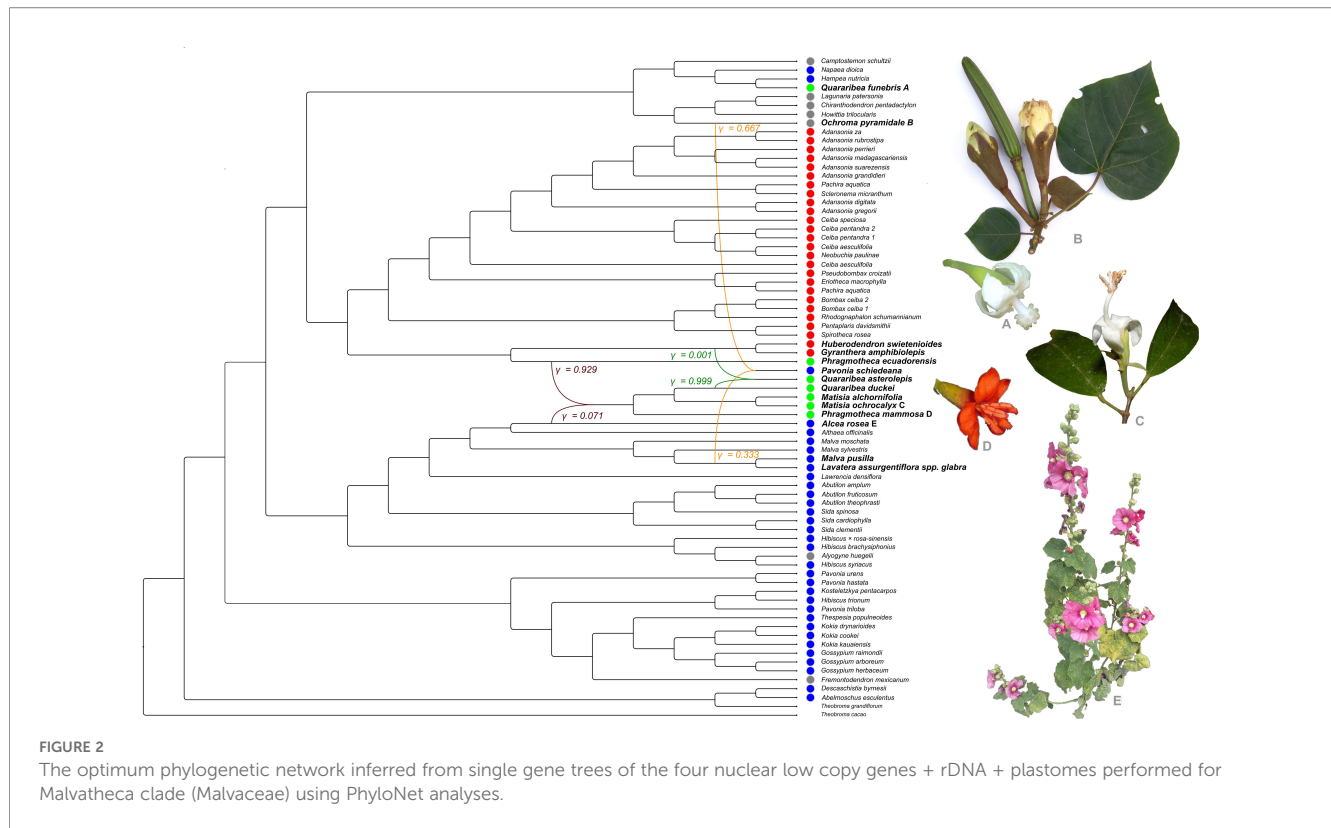
Clade IV is morphologically unclear composed of three *incertae sedis* species (*Chiranthodendron pentadactylon*, *Lagunaria patersoniana* and *Ochroma pyramidale*), one Matisioideae (*Q. funebris*), and two Malvoideae species (*Hampea nutricia* and *Napea dioica*). Additionally *Pavonia schiedeana* was an exception, appearing in this clade only in the low-copy nuclear and rDNA topologies, not in the plastid topology (Figure 1). In the low copy nuclear genes topology, these species were placed within the Clade II. The incongruence and relationship details are present in Supplementary Figures 1, 2.

Reticulate evolution in Malvatheca

By assuming positional changes in nuclear and plastidial topologies as evidence of reticulate evolution, our data revealed a high complexity among plastidial, rDNA, and low-copy nuclear gene topologies (Figure 1). Major topological incongruences involved the placement of two *Pavonia* species. *P. urens* was resolved within Clade III in the plastid topology, while *P. schiedeana* was placed in Clade IV in both the low-copy nuclear and rDNA topologies. The PhyloNet analysis provided strong evidence for reticulate evolution in Malvatheca. The DeepCoalCount function performed a species tree with a 'total number of extra lineages score' of 695.0. Subsequent testing the distinct reticulation hypotheses (H_1 – H_4) using the DeepCoalCount_network algorithm yielded the following scores: $H_1 = 517.0$, $H_2 = 497.0$, $H_3 = 456.0$, and $H_4 = 466.0$ (The lower this score, the less conflict there is between the tested loci). Thus, H_3 was selected and this network revealed three statistically supported reticulation events within the Malvatheca clade (Figure 2): (i) A hybrid origin of *Pavonia schiedeana* Steud. involving the cross between the lineages of *Malva pusilla* Sm. + *Lavatera assurgentiflora* subsp. *glabra* (Cockerell) Guillems clade (inheritance proportion, $\gamma = 0.333$) and *Ochroma pyramidale* (Cav. ex Lam.) Urb. ($\gamma = 0.667$); (ii) a hybrid origin of *Quararibea asterolepis* Pittier by the cross between the lineages *Huberodendron swietenoides* (Gleason) Ducke + *Gyranthra amphibiolepis* W.Palacios clade ($\gamma = 0.001$) and *Quararibea duckei* Huber ($\gamma = 0.999$); (iii) a ancient hybridization between ancestors lineages of *Phragmothea ecuadorensis* W.S.Alverson ($\gamma = 0.929$) and *Alcea*



(*Quararibea*) and the genus *Huberodendron* ($D = 0.54$), confirming that the observed topological discordance is not merely the result of incomplete lineage sorting (ILS), but rather of significant ancient gene flow. On the other hand, the reticulation at the base of the Matisoioideae clade showed $D = 0.05$, with the distribution of shared (ABBA) and unshared (BABA) alleles being almost perfectly symmetrical (19 to 17), suggesting a scenario of Incomplete Lineage Sorting (ILS).



Repeatome diversity in the Malvatheca clade

In the individual repeatome analyses, the sequencing depth ranged from 40,572 reads in *Bombax ceiba* (accession 2) to 2,633,822 in *Hibiscus syriacus* (Supplementary Table 1). The proportion of repetitive DNA varied widely, from 4.8% in *Pseudobombax croizatii* A. Robyns to 56.4% in *Spirotheca rosea* (Seem.) P.E.Gibbs & W.S. Alverson (Bombacoideae), and from 3.4% in *Howittia trilocularis* to 65.4% in *Abutilon amplum* Benth. (Malvoideae) (Figure 3; Supplementary Table 2).

Both subfamilies were dominated by Class I transposable elements, particularly LTR retrotransposons of the *Ty3-type* superfamily. Their abundance ranged from 0.5% in *Malva sylvestris* L. to 50% in *Chiranthodendron pentadactylon* Larreat. (Supplementary Table 2). Within Bombacoideae, Tekay elements predominated across four genera, whereas Ogre elements also occurred at moderate to high levels. In Malvoideae, Tekay and Ogre were generally present in moderate proportions, with a few species showing elevated abundances of Athila elements. Among *Ty1-copia* elements, Angela and SIRE were the most widespread across both subfamilies, with SIRE being particularly enriched in *Sida cardiophylla* Domin. and one *Bombax ceiba* accession (Figure 3; Supplementary Table 3). The satellite DNA content also varied considerably, from 0.02% in *Gossypium arboreum* L. to 33.9% in *Adansonia grandidieri* Baill. Bombacoideae, especially *Adansonia*, tended to have relatively

high satellite DNA abundances (4.9–33.8%). Despite this variation, no genomic synapomorphies based on repeat presence/absence were detected that clearly distinguished Bombacoideae from Malvoideae (Figure 3; Supplementary Tables 2, 3).

Comparative repeatome analysis revealed limited sharing of repetitive DNA clusters between subfamilies. Within Bombacoideae, intragenomic similarity was high: *Adansonia* species shared the most repeat clusters, and *Pseudobombax croizatii* and *Pachira aquatica* Aubl. Both genera presented elevated proportions of *Ty3-type* Ogre and CRMs. In contrast, Malvoideae species presented highly divergent repeatomes, with even congeneric taxa (e.g., *Malva sylvestris* and *M. moschata* L.) sharing few clusters (Supplementary Figure 3; Supplementary Table 4).

The repeat-based phylogeny recovered three major clades but with generally weak support (Supplementary Figure 4; Supplementary Table 4). Clade III showed extensive mixing, grouping representatives of all subfamilies without resolution. Several genera were not recovered as monophyletic, including *Adansonia* (Bombacoideae) and *Abutilon*, *Sida*, *Hibiscus*, and *Malva* (Malvoideae). Three Bombacoideae species (*Phragmothecha mammosa*, *Gyranthra amphibiolepis* W. Palacios, *Huberodendron swietenoides* (Gleason) Ducke) and four Malvoideae species (*Camptostemon schultzei* Mast., *Pentaplaris davidsmithii* Dorr & C. Bayer, *Alyogyne huegelii* (Endl.) Fryxell, *Pavonia schiedeana* Steud.) shifted into clade III. Additional rearrangements, such as altered relationships among *Adansonia*, *Scleronema* Benth, and

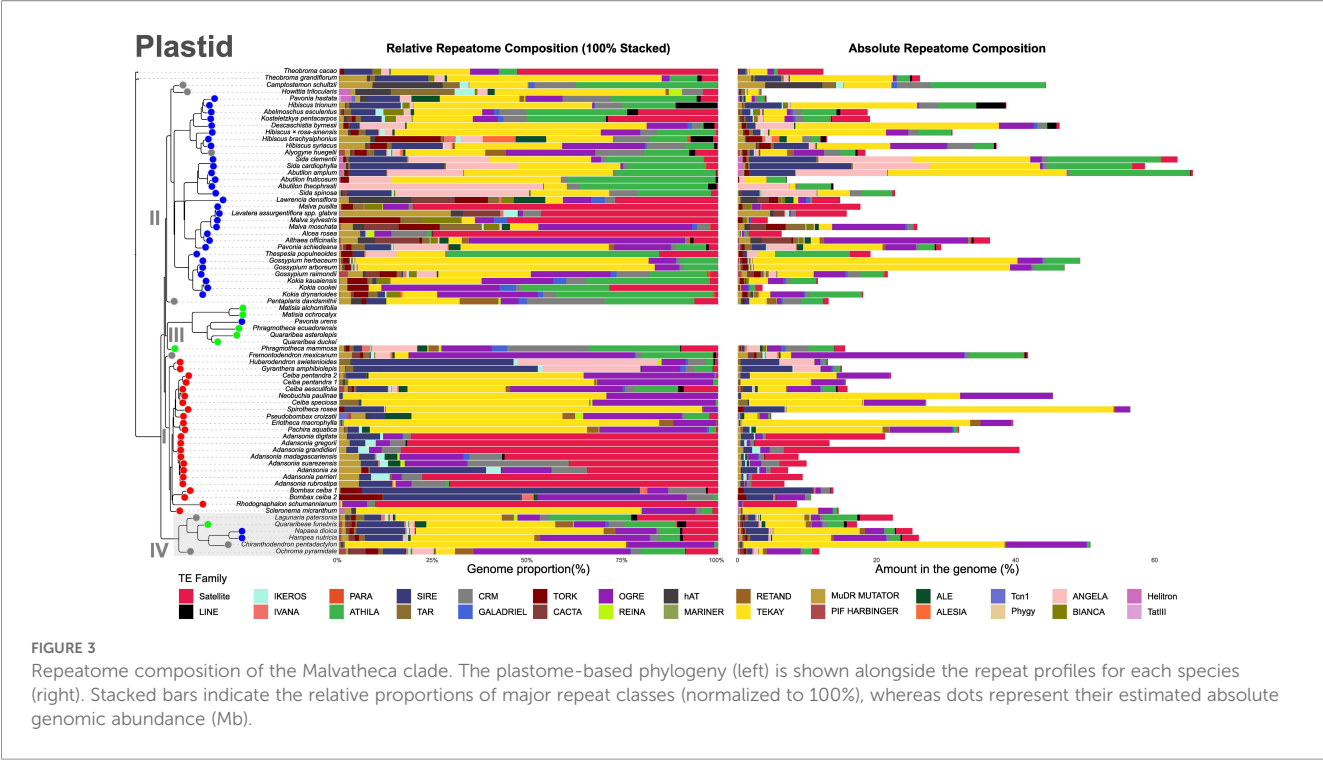


FIGURE 3 Repeatome composition of the Malvatheca clade. The plastome-based phylogeny (left) is shown alongside the repeat profiles for each species (right). Stacked bars indicate the relative proportions of major repeat classes (normalized to 100%), whereas dots represent their estimated absolute genomic abundance (Mb).

Rhodognaphalon (Ulbr.) Roberty (clade I) and the placement of *Pavonia*, *Decaschistia*, and two *Hibiscus* species in clade II (Supplementary Figure 4; Supplementary Table 4).

Comparison with the rDNA tree revealed similar patterns: multiple species from clades I and II were grouped into clade III, with further positional changes within clades (e.g., *Bombax ceiba*, *Fremontodendron*, and *Rhodognaphalon* in clade I; *Howittia*, *Hibiscus*, and *Decaschistia byrnesii* Fryxell in clade II) (Supplementary Figure 4; Supplementary Table 4).

Diversity of repeats and correlation with chromosome number

The repeat diversity of Malvatheca, as measured by the Shannon (H) and Simpson (D) indices, ranged from very low values in *Rhodognaphalon schumannianum* (A. Robyns) A. Robyns (H = 0.4, D = 0.17) to high values in *Hibiscus brachysiphonius* F. Muell. (H = 2.4, D = 0.9). Malvoideae species generally presented greater repeat diversity (H = 0.8–2.4; D = 0.3–0.9) than did Bombacoideae, which tended toward lower values (Figure 4a; Supplementary Table 5). Chromosome number was negatively correlated with repeat diversity (Shannon: R = −0.36, p = 0.022; Simpson: R = −0.39, p = 0.011), indicating that species with higher chromosome counts, such as Bombacoideae, typically presented fewer diverse repeatomes, whereas Malvoideae (lower 2n) presented greater diversity (Figure 4b; Supplementary Table 5).

To assess whether chromosome number is correlated with growth habits, we compiled data from 84 Malvaceae species. Trees presented consistently greater chromosome numbers than shrubs and herbs did (range: 8–328; median = 30), supporting an

association between the arboreal habit and large chromosome complements (Figure 4b; Supplementary Table 6).

Discussion

Would the diversification of Malvatheca be driven by reticulate evolution?

Our data provides phylogenetic evidence for RE in the Malvatheca clade, corroborating previous genomic analyses (Karimi et al., 2020; Wan et al., 2024; Zhang et al., 2025). These reticulation events appear to have occurred on different timescales (Hernández-Gutiérrez et al., 2022; Yang et al., 2025; Stull et al., 2023), ranging from more recent lineages (e.g., the 21.6 Mya genus *Adansonia*; Wan et al., 2024) to the 53.5 Mya origin of Bombacoideae (Zizka et al., 2020). Hybridization has long been recognized as a fundamental evolutionary force shaping plant diversity, with impacts ranging from immediate reproductive isolation to long-term adaptive radiation (Rieseberg, 1997; Soltis et al., 2004; Chester et al., 2012). Phylogenetic analysis of Malvaceae revealed a complex reticulated evolutionary history. Network inference (PhyloNet) under Maximum Parsimony and Maximum Likelihood criteria rejected strict bifurcation in favor of a network with three hybridization events. The clades III and IV were consistently associated with reticulation points in both plastid and nuclear datasets, suggesting a complex hybrid origin (Hernández-Gutiérrez et al., 2022; Yang et al., 2025). The morphological and cytogenetic heterogeneity of its members supports this view. Sampling the only three genera of Matisioideae (*Matisia*, *Quararibea*, and *Phragmothea*) allowed us to demonstrate that

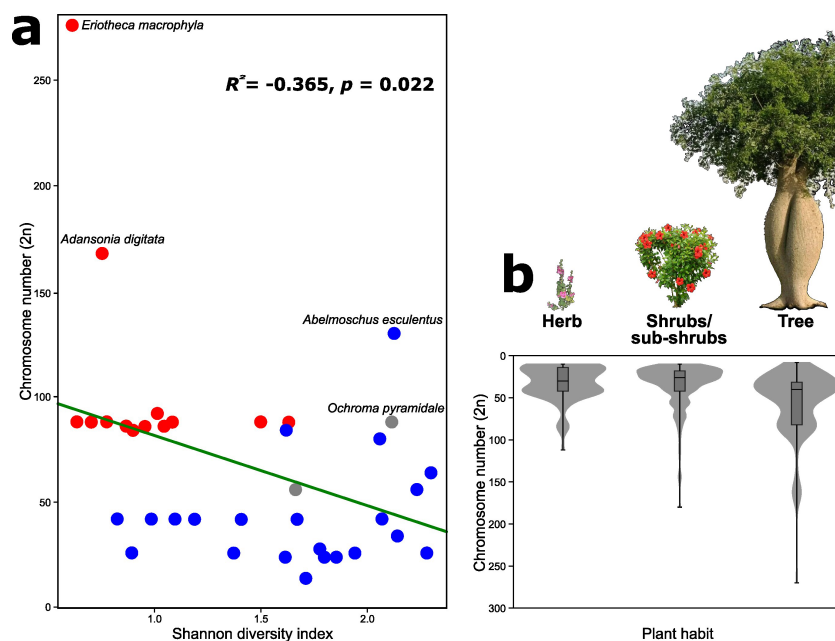


FIGURE 4

Repeat diversity and chromosome number variation in Malvaceae. (a) Shannon diversity index of repetitive DNA elements in Malvaceae species, colored by taxonomic group: Bombacoideae (red), Malvoideae (blue), and incertae sedis (gray). (b) Chromosome number distributions across growth habits (herbs, shrubs/subshrubs, and trees) compiled from Malvaceae species.

most representatives of the subfamily (except *Q. funebris*) are grouped into a well-supported clade. The hybrid origin test of this clade, involving a reticulation event between Bombacoideae and Malvoideae, was inconclusive, and our analyses suggest that it may simply be a case of high levels of deep incomplete lineage sorting (ILS) among these subfamilies. Regarding the methods used here to calculate hybridization or introgression (D-statistics), it is worth noting that these tests can be highly sensitive to evolutionary rate variation across different lineages, which can frequently lead to false positives or inconclusive results (Frankel and Ané, 2023). Given the phylogenetic, morphological, and systematic complexity of the relationship between the Malvaceae subfamilies (Colli-Silva et al., 2025), these hypotheses of reticulated evolution need to be tested further with a larger sample size. There are no widely sampled phylogenetic hypotheses for the genus *Quararibea*; therefore, the unexpected position of *Q. funebris* would need to be investigated in more detail, and may be related to taxonomic identification error, reticulation, and/or methodological limitations (e.g., low coverage).

While nuclear low copy genes, rDNA and plastid datasets provided strong phylogenetic resolution for Malvaceae, the repeat-based approach (Dodsworth et al., 2015) failed to resolve relationships. This loss of signal may reflect (i) the impact of reticulation, as hybridization, particularly in allopolyploid contexts, can induce substantial restructuring of repetitive elements through sequence loss, differential retention of parental repeats, and epigenetic reorganization, thereby altering repeat presence, absence, and relative abundance (Parisod et al., 2010), and/or (ii) the deep divergence of Bombacoideae (53.5 Mya; Zizka et al., 2020). Because repeats undergo rapid evolutionary turnover, their phylogenetic utility can erode in ancient groups. Nevertheless,

recent studies have shown that repeatome composition can retain signals and even capture ancient hybridization through characteristic proliferation patterns (Castro et al., 2024; Hlavatá et al., 2024). The contrast across plant groups highlights the lineage- and timescale-specific nature of repeat data. For example, in *Erythrostemon* Klotzsch (Leguminosae, 33.6 Mya), systematic variation in repeat composition revealed discordant profiles consistent with ancient origins (Castro et al., 2024). Similarly, in *Amomum* L. (Zingiberaceae, 19.3 Mya), bursts of repeat proliferation coincided with inferred hybridization events, driving genome size changes that track clade divergence (Li et al., 2023; Hlavatá et al., 2024).

Genomic-scale phylogenies further demonstrated that reticulate evolution is pervasive across Malvaceae, challenging tree-based models (Hernández-Gutiérrez et al., 2022; Yang et al., 2025; Zhang et al., 2025). Using 268 nuclear loci across 96 genera, Hernández-Gutiérrez and Magallon (2019) detected extensive discordance attributable to incomplete lineage sorting and introgression, showing that bifurcating trees fail to capture the family's evolutionary complexity. Reticulation manifests at multiple scales, from recent interspecific gene flow to ancient allopolyploid events. Hyb-Seq datasets and network inference have revealed well-supported introgression in baobabs (*Adansonia*), clarifying floral homoplasmy and pollination biology (Karimi et al., 2020). Similarly, independent hybridization and introgression across *Gossypium* highlight the role of allopolyploidy in diversification (Zhang et al., 2025). Cytonuclear processes in *Hibiscus* (Pfeil, 2002) involve chloroplast capture, whereas cytogenetic work in *Eriotheca* (Serra et al., 2022) confirms that hybridization and polyploidy remain active processes. Together, these findings underscore that

phylogenetic networks, rather than bifurcated trees, more accurately represent Malvaceae evolution and highlight the central role of RE in shaping its genomic and phenotypic diversity.

The repeatome compositions of the Malvatheca clade subfamilies

Genomic studies indicate that the Malvatheca clade has undergone multiple rounds of polyploidization, including reticulate allopolyploidy and subsequent dysploidy, which have played central roles in its diversification (Hernández-Gutiérrez et al., 2022; Yang et al., 2025; Zhang et al., 2025). Recent chromosome-scale assemblies of *Adansonia* spp., *Bombax ceiba* and *Ceiba pentandra* have provided unprecedented insights into genome structure and repeat composition across Malvaceae (Wan et al., 2024; Shao et al., 2024). Complementary phylogenomic studies, such as those on *Hibiscus* L., have revealed at least three independent whole-genome duplication (WGD) events (Eriksson et al., 2021), highlighting the recurrent nature of polyploidy within the family. These duplications supplied raw genetic material for evolutionary novelty, whereas subsequent post-polyploid diploidization processes reshaped karyotypes and contributed to current genomic diversity patterns.

The broader evolutionary significance of WGD has been well documented across angiosperms. Large-scale comparative studies have shown that WGDs are not randomly distributed in time but cluster during episodes of environmental stress (Van de Peer et al., 2017; Teng et al., 2022). This has given rise to the so-called “polyploidy paradox”: although WGD frequently occurs, only a subset of polyploid lineages survive in the long term (Van de Peer et al., 2021). Several of these surviving events appear to coincide with the Cretaceous–Paleogene (K–Pg) boundary (~66 Mya), a period of mass extinction. WGDs at this time are hypothesized to have facilitated lineage survival and diversification by buffering against genomic stress and providing redundancy for adaptive innovation (Fawcett et al., 2009; Baduel et al., 2018; Bomblies, 2020). The “polyploid hop” model further proposes that polyploidy creates both challenges and opportunities: initial barriers to establishment may be overcome by long-term adaptive potential (Baduel et al., 2018). Recent studies also suggest that specific life-history traits, such as seed biology, may interact with polyploid status to influence which lineages persist across the K–Pg boundary (Cai et al., 2019; Berry and Jaganathan, 2022).

Within Malvaceae, repeatome dynamics provide important clues to post-WGD genome evolution. Chromosome-scale assemblies have shown that transposable element (TE) proliferation and chromosome restructuring are major forces shaping genome size and karyotype variation (Li et al., 2024; Shao et al., 2024). In Malvatheca, distinct repeatome profiles are observed among subfamilies: Bombacoideae tends to exhibit lower repeat diversity, Malvoideae has higher abundance and diversity, and Matisioideae has intermediate patterns. These differences likely reflect contrasting evolutionary strategies for restructuring genomes after polyploidy, with Bombacoideae favoring karyotypic stability through TE

amplification, Malvoideae accumulating a more diverse repeat fraction, and Matisioideae retaining a mixed profile.

Taken together, these patterns suggest that the interplay among ancient WGDs, repeatome evolution, and post-polyploid diploidization has been a major driver of lineage-specific diversification in Malvatheca. Variation in repeat composition and chromosome number across subfamilies reflects different genomic strategies for managing the legacy of polyploidy, highlighting the central role of repetitive DNA in shaping both the stability and flexibility of plant genomes over deep evolutionary timescales.

The habit correlated with chromosome number

Our results revealed that tree-dominated lineages presented the highest chromosome numbers. In Bombacoideae, striking examples include *Pseudobombax* and *Pachira* ($2n = 88–92$), *Adansonia digitata* ($2n = 160$), and *Eriotheca* species, with exceptionally elevated counts of $2n = 194–276$ (Costa et al., 2017). A similar pattern was described for incertae sedis, such as *Ochroma pyramidale* ($2n = 84$) (Costa et al., 2017). A similar trend occurs in the exclusively arboreal Tilioideae, where counts range from $2n = 82$ to $2n = 328$ (Pigott, 2002; 2012). These data reflect the trend toward ancient whole-genome duplication (WGD) events, followed by the lineage-specific conservation of high chromosome numbers in long-life-cycle species.

Shrubby lineages, such as *Hibiscus* ($2n \approx 52$) and *Pavonia* ($2n \approx 56$), present intermediate values, possibly representing transitional stages between herbaceous and woody forms. In contrast, herbaceous genera such as *Sida* are chromosomally conserved, typically ranging from $2n = 14–32$ (Skovsted, 1935; Fernández et al., 2003). These reductions likely result from dysploidy — the step-wise loss or fusion of chromosomes — a process often associated with short-lived, fast-reproducing life cycles (Siljak-Yakovlev et al., 2017).

The habit–chromosome relationship thus reflects contrasting evolutionary pressures. In woody clades such as Bombacoideae, ancient polyploidization established high baselines that were further expanded by more recent polyploidy (Marinho et al., 2014). Herbaceous lineages, in turn, underwent systematic reductions, which is consistent with patterns observed in other families, such as Asteraceae, where perennial-to-annual transitions are linked to smaller karyotypes (Watanabe et al., 1999). Selective pressures on herbaceous species — favoring rapid cell cycles and short generation times — likely reinforced this reduction. Similar correlations between life form and karyotype have been documented across angiosperms (e.g., Malpighiaceae, Lombello and Forni-Martins, 2003; Ehrendorfer, 1989), but Malvaceae provides one of the clearest examples: Bombacoideae, dominated by trees, consistently carries higher chromosome numbers than Malvoideae, which include mostly herbs and shrubs. These results suggest that major life form transitions are accompanied by fundamental changes in genome organization.

Conclusions

Our phylogenomic analyses provide strong evidence for reticulate evolution in the Malvaceae clade (Malvaceae), underscoring its role as a fundamental driver of diversification in the family. Ancient hybridization events, coupled with allopolyploidy and subsequent diploidization, have left lasting imprints on the genomic architecture and subfamily relationships. Nuclear and plastid datasets revealed extensive reticulation, whereas repeatome analyses proved less effective for deep phylogenetic resolution, reflecting rapid lineage-specific turnover. This contrast highlights the importance of multilayered genomic approaches for reconstructing complex evolutionary histories. Comparative genomic and cytogenetic evidence also reveals subfamily specific patterns in karyotype evolution. Tree-dominated lineages, particularly Bombacoideae, exhibit exceptionally high and variable chromosome numbers arising from successive polyploidization, whereas herbaceous taxa display reduced counts through dysploidy and genome downsizing. These correlations suggest a close interplay between life form, polyploidy, and genome structure: woody lineages preserve high baselines established by ancient WGDs, whereas herbaceous taxa evolve toward reduced karyotypes under different life-history constraints. Taken together, our results demonstrate how hybridization, polyploidy, and repeat dynamics have jointly shaped Malvaceae diversification. This family offers a compelling model for understanding the integration of reticulate processes, structural genome evolution, and life-history strategies in driving morphological and ecological innovation. Future research incorporating chromosome-scale assemblies and functional studies will be critical for disentangling the genomic consequences of WGD and reticulation across this diverse plant lineage.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Author contributions

GL: Formal Analysis, Investigation, Methodology, Visualization, Writing – original draft. LC: Data curation, Formal Analysis, Investigation, Writing – review & editing. FP: Data curation, Writing – review & editing. NK: Data curation, Writing – review & editing. JS: Data curation, Writing – review & editing. JC-S: Data curation, Writing – review & editing. MC-S: Data curation, Writing – review & editing. AM: Funding acquisition, Resources, Writing –

original draft, Writing – review & editing. GS: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq number Universal 407535/2023-3). Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for GL MSc grant (process number 88887.959545/2024-00). GS received a productivity fellowship from CNPq (process number PQ 305774/2024-7).

Acknowledgments

The authors thank the Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq number Universal 407535/2023-3) for financial support. GL thanks the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (Process number 88887.959545/2024-00) for the MSc. grant. A.M. is financially supported by the Max Planck Society. This study was financed in part by the CAPES (Brazil)—Finance Code 001. G. received a productivity fellowship from CNPq (process number PQ 305774/2024-7). We also thank the ELIXIR- CZ Research Infrastructure Project (LM2015047) for providing computational resources for RepeatExplorer analysis.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. During the preparation of this work, the author(s) used ChatGPT in order to check grammatical errors. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2025.1717745/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Circular tanglegram comparing phylogenetic topologies inferred from low-copy nuclear genes (Genes), plastid genomes (Plastid), and ribosomal DNA (rDNA) for *Malvatheca* (Malvaceae). Branches and connecting links are colored according to taxonomic groups: Bombacoideae (red), Malvoideae

(blue), Matisioideae (green), and *Incertae sedis* (grey). The central lines connect corresponding taxa between adjacent trees to visualize topological congruence and conflict. Roman numerals (I–IV) indicate major clades recovered in the analyses.

SUPPLEMENTARY FIGURE 2

Phylogenetic relationships inferred from whole rDNA, plastome, low copy nuclear genes and repeat abundance data. Maximum Likelihood trees for the rDNA, plastome and low copy nuclear genes are shown followed by a Neighbor-Joining tree based on repeat abundance. Support values are shown on nodes for all trees.

SUPPLEMENTARY FIGURE 3

Phylogeny and repeatome analysis of the *Malvatheca* clade (Malvaceae). The plastome-based phylogeny (left) shows the evolutionary relationships between species. The heatmap (right) compares the abundance of different repetitive element classes across these species, with brown indicating high abundance and light colors indicating low abundance.

SUPPLEMENTARY FIGURE 4

Phylogenetic relationships within the *Malvatheca* clade inferred from inferred from plastome, nuclear rDNA and low copy nuclear genes (left) and repeats abundance (right) sequence data. Tip labels are color-coded by taxonomic group: Malvoideae (blue), Bombacoideae (red), Matisioideae (green), and *incertae sedis* (gray).

References

- Arnold, M. L. (1997). *Natural Hybridization and Evolution* (Oxford: Oxford University Press).
- Baduel, P., Bray, S., Vallejo-Marín, M., Kolář, F., and Yant, L. (2018). The “Polyploid Hop”: shifting challenges and opportunities over the evolutionary lifespan of genome duplications. *Front. Ecol. Evol.* 6, 117. doi: 10.3389/fevo.2018.00117
- Baum, D. A., DeWitt Smith, S., Yen, A., Alverson, W. S., Nyffeler, R., Whitlock, B. A., et al. (2004). Phylogenetic relationships of *Malvatheca* (Bombacoideae and Malvoideae; Malvaceae sensu lato) as inferred from plastid DNA sequences. *Am. J. Bot.* 91, 1863–1871. doi: 10.3732/ajb.91.11.1863
- Berry, K., and Jaganathan, G. K. (2022). Did selection for seed traits across the Cretaceous/Paleogene boundary sort plants based on ploidy? *Acta Palaeobot.* 62, 182–195. doi: 10.35535/acpa-2022-0012
- Bomblies, K. (2020). When everything changes at once: finding a new normal after genome duplication. *Proc. R. Soc. B.* 287, 20202154. doi: 10.1098/rspb.2020.2154
- Bowers, J. E., and Paterson, A. H. (2021). Chromosome number is key to longevity of polyploid lineages. *New Phytol.* 231, 19–28. doi: 10.1111/nph.17361
- Cai, L., Xi, Z., Amorim, A. M., Sugumaran, M., Rest, J. S., Liu, L., et al. (2019). Widespread ancient whole-genome duplications in Malpighiales coincide with Eocene global climatic upheaval. *New Phytol.* 221, 565–576. doi: 10.1111/nph.15357
- Carvalho-Sobrinho, J. G., Alverson, W. S., Alcantara, S., Queiroz, L. P., Mota, A. C., and Baum, D. A. (2016). Revisiting the phylogeny of Bombacoideae (Malvaceae): Novel relationships, morphologically cohesive clades, and a new tribal classification based on multilocus phylogenetic analyses. *Mol. Phylogenet. Evol.* 101, 56–74. doi: 10.1016/j.ympev.2016.05.006
- Castro, N., Vilela, B., Mata-Sucre, Y., Marques, A., Gagnon, E., Lewis, G. P., et al. (2024). Repeatome evolution across space and time: Unravelling repeats dynamics in the plant genus *Erythrostemon* Klotzsch (Leguminosae Juss). *Mol. Ecol.* 33, e17510. doi: 10.22541/au.172457451.19887838/v1
- Cavallini, A., Mascagni, F., Giordani, T., and Natali, L. (2019). Genome skimming for plant retrotransposon identification and expression analysis. *Agrochimica* 63, 367–378. doi: 10.12871/00021857201945
- Chester, M., Gallagher, J. P., Symonds, V. V., Cruz da Silva, A. V., Mavrodiev, E. V., Leitch, A. R., et al. (2012). Extensive chromosomal variation in a recently formed natural allopolyploid species, *Tragopogon miscellus* (Asteraceae). *Proc. Natl. Acad. Sci. U.S.A.* 109, 1176–1181. doi: 10.1073/pnas.1112041109
- Colli-Silva, M., Pérez-Escobar, O. A., Ferreira, C. D., Costa, M. T., Gerace, S., Coutinho, T. S., et al. (2025). Taxonomy in the light of incongruence: An updated classification of Malvales and Malvaceae based on phylogenomic data. *Taxon* 74, 361–385. doi: 10.1002/tax.13300
- Costa, L., Marques, A., Buddenhagen, C., Thomas, W. W., Huettel, B., Schubert, V., et al. (2021). Aiming off the target: recycling target capture sequencing reads for investigating repetitive DNA. *Ann. Bot.* 128, 835–848. doi: 10.1093/aob/mcab063
- Costa, L., Oliveira, Á., Carvalho-Sobrinho, J., and Souza, G. (2017). Comparative cytomic analyses reveal karyotype variability related to biogeographic and species richness patterns in Bombacoideae (Malvaceae). *Plant Syst. Evol.* 303, 1131–1144. doi: 10.1007/s00606-017-1427-6
- Cvetković, T., Areces-Berzain, F., Hinsinger, D. D., Thomas, D. C., Wieringa, J. J., Ganesan, S. K., et al. (2021). Phylogenomics resolves deep subfamilial relationships in Malvaceae s.l. G3 11, jkab136. doi: 10.1093/g3journal/jkab136
- Dodsworth, S. (2015). Genome skimming for next-generation biodiversity analysis. *Trends Plant Sci.* 20, 525–527. doi: 10.1016/j.tplants.2015.06.012
- Dodsworth, S., Chase, M. W., Kelly, L. J., Leitch, I. J., Macas, J., Novak, P., et al. (2015). Genomic repeat abundances contain phylogenetic signal. *Syst. Biol.* 64, 112–126. doi: 10.1093/sysbio/syu080
- Dodsworth, S., Pokorny, L., Johnson, M. G., Kim, J. T., Maurin, O., Wickett, N. J., et al. (2019). Hyb-Seq for flowering plant systematics. *Trends Plant Sci.* 24, 887–891. doi: 10.1016/j.tplants.2019.07.011
- Durand, E. Y., Patterson, N., Reich, D., and Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* 28, 2239–2252. doi: 10.1093/molbev/msr048
- Ehrendorfer, F. (1989). Progrès des connaissances sur la différenciation caryologique et sur l'évolution des plantes ligneuses tropicales. *Bull. Soc. Bot. Fr. Actual.* Bot. 136, 197–206. doi: 10.1080/01811789.1989.10826974
- Eriksson, J. S., Bacon, C. D., Bennett, D. J., Pfeil, B. E., Oxelman, B., and Antonelli, A. (2021). Gene count from target sequence capture places three whole genome duplication events in *Hibiscus* L. (Malvaceae). *BMC Ecol. Evol.* 21, 107. doi: 10.1186/s12862-021-01751-7
- Fawcett, J. A., Maere, S., and Van De Peer, Y. (2009). Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5737–5742. doi: 10.1073/pnas.0900906106
- Fawcett, J. A., and Van de Peer, Y. (2010). Angiosperm polyploids and their road to evolutionary success. *Trends Evol. Biol.* 2, e3. doi: 10.4081/eb.2010.e3
- Fernández, A., Krapovickas, A., Lavia, G., and Seijo, G. (2003). Cromosomas de malváceas. *Bonplandia* 12, 141–145. doi: 10.30972/bon.121-41411
- Frankel, L. E., and Ané, C. (2023). Summary tests of introgression are highly sensitive to rate variation across lineages. *Systematic Biol.* 72, 1357–1369. doi: 10.1093/sysbio/syad056
- Gontier, N. (2015). “Reticulate evolution everywhere,” in *Reticulate Evolution: Symbiogenesis, Lateral Gene Transfer, Hybridization and Infectious Heredity*. Ed. N. Gontier (Springer International Publishing, Cham), 1–40.
- Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., et al. (2010). A draft sequence of the Neandertal genome. *Science* 328, 710–722. doi: 10.1126/science.1188021

- Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize Implements and enhances circular visualization in R. *Bioinformatics* 30, 2811–2812. doi: 10.1093/bioinformatics/btu393
- Hammer, Ø., Harper, D. A., and Ryan, P. D. (2001). PAST: paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* 4, 9.
- Hernández-Gutiérrez, R., and Magallon, S. (2019). The timing of Malvales evolution: Incorporating its extensive fossil record to inform about lineage diversification. *Mol. Phylogenet. Evol.* 140, 106606. doi: 10.1016/j.ympev.2019.106606
- Hernández-Gutiérrez, R., van den Berg, C., Granados Mendoza, C., Peñafiel Cevallos, M., Freire, M. E., Lemmon, E. M., et al. (2022). Localized phylogenetic discordance among nuclear loci due to incomplete lineage sorting and introgression in the family of cotton and cacao (Malvaceae). *Front. Plant Sci.* 13, 850521. doi: 10.3389/fpls.2022.850521
- Hlavatá, K., Závěská, E., Leong-Škorničková, J., Pouch, M., Poulsen, A. D., Šída, O., et al. (2024). Ancient hybridization and repetitive element proliferation in the evolutionary history of the monocot genus *Amomum* (Zingiberaceae). *Front. Plant Sci.* 15, 1324358. doi: 10.3389/fpls.2024.1324358
- Huson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267. doi: 10.1093/molbev/msj030
- Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., et al. (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature* 473, 97–100. doi: 10.1038/nature09916
- Karstein, K., Tomasello, S., Hodač, L., Wagner, N., Marínček, P., Barke, B. H., et al. (2022). Untying Gordian knots: unraveling reticulate polyploid plant evolution by genomic data using the large *Ranunculus auricomus* species complex. *New Phytologist* 235:2081–2098.
- Karimi, N., Grover, C. E., Gallagher, J. P., Wendel, J. F., Ané, C., and Baum, D. A. (2020). Reticulate evolution helps explain apparent homoplasy in floral biology and pollination in baobabs (*Adansonia*; Bombacoideae; Malvaceae). *Syst. Biol.* 69, 462–478. doi: 10.1093/sysbio/sy073
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
- Li, D. M., Liu, H. L., Pan, Y. G., Yu, B., Huang, D., and Zhu, G. F. (2023). Comparative chloroplast genomics of 21 species in Zingiberales with implications for their phylogenetic relationships and molecular dating. *Int. J. Mol. Sci.* 24, 15031. doi: 10.3390/ijms241915031
- Li, W., Chen, X., Yu, J., and Zhu, Y. (2024). Upgraded durian genome reveals the role of chromosome reshuffling during ancestral karyotype evolution, lignin biosynthesis regulation, and stress tolerance. *Sci. China Life Sci.* 67, 1266–1279. doi: 10.1007/s11427-024-2580-3
- Lombello, R. A., and Forni-Martins, E. R. (2003). Malpighiaceae: correlations between habit, fruit type and basic chromosome number. *Acta Bot. Bras.* 17, 171–178. doi: 10.1590/S0102-33062003000200001
- Mallet, J. (2007). Hybrid speciation. *Nature* 446, 279–283. doi: 10.1038/nature05706
- Marinho, R. C., Mendes-Rodrigues, C., Balao, F., Ortiz, P. L., Yamagishi-Costa, J., Bonetti, A. M., et al. (2014). Do chromosome numbers reflect phylogeny? New counts for Bombacoideae and a review of Malvaceae s.l. *Am. J. Bot.* 101, 1456–1465. doi: 10.3732/ajb.1400248
- Novák, P., Neumann, P., and Macas, J. (2020). Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. *Nat. Protoc.* 15, 11. doi: 10.1038/s41596-020-0400-y
- Parisod, C., Alix, K., Just, J., Petit, M., Sarilar, V., Mhiri, C., et al. (2010). Impact of transposable elements on the organization and function of allopolyploid genomes. *New Phytol.* 186, 37–45. doi: 10.1111/j.1469-8137.2009.03096.x
- Paterson, A. H., Bowers, J. E., and Chapman, B. (2004). Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9903–9908. doi: 10.1073/pnas.0307901101
- Pezzi, F. F., Dexter, K. G., de Carvalho-Sobrinho, J. G., Kidner, C. A., Nicholls, J. A., de Queiroz, L. P., et al. (2021). Phylogeny and biogeography of Ceiba Mill. (Malvaceae, Bombacoideae). *Front. Biogeogr.* 13. doi: 10.21425/F5FBG49226
- Pfeil, B. E. (2002). *Phylogenetic systematics of Hibiscus and related genera (Malvaceae)* (Ithaca: Cornell University).
- Pigott, C. D. (2002). A review of chromosome numbers in the genus *Tilia* (Tiliaceae). *Edinb. J. Bot.* 59, 239–246. doi: 10.1017/S0960428602000057
- Pigott, D. (2012). *Lime-trees and basswoods: a biological monograph of the genus Tilia* (Cambridge: Cambridge University Press).
- R Core Team (2019). *R: A Language and Environment for Statistical Computing* (Vienna: R Foundation for Statistical Computing).
- Rieseberg, L. H. (1997). Hybrid origins of plant species. *Annu. Rev. Ecol. Syst.* 28, 359–389. doi: 10.1146/annurev.ecolsys.28.1.359
- Schley, R. J., Pellicer, J., Ge, X. J., Barrett, C., Bellot, S., Guignard, M. S., et al. (2022). The ecology of palm genomes: Repeat-associated genome size expansion is constrained by aridity. *New Phytol.* 236, 433–446. doi: 10.1111/nph.18323
- Serra, A. D., Rodrigues, C. M., Marinho, R. C., Balao, F., and Oliveira, P. E. (2022). Interploidy hybridization in *Eriotheca gracilipes* and *E. pubescens* (Malvaceae): experimental evidence, genome and stomatal size. *Rodriguésia* 73, e01292021. doi: 10.1590/2175-7860202273078
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x
- Shao, L., Jin, S., Chen, J., Yang, G., Fan, R., Zhang, Z., et al. (2024). High-quality genomes of Bombax ceiba and Ceiba pentandra provide insights into the evolution of Malvaceae species and differences in their natural fiber development. *Plant Commun.* 5, 100784. doi: 10.1016/j.xplc.2024.100832
- Siljak-Yakovlev, S., Godelle, B., Zoldos, V., Vallès, J., Garnatje, T., and Hidalgo, O. (2017). Evolutionary implications of heterochromatin and rDNA in chromosome number and genome size changes during dysploidy: A case study in Reichardia genus. *PLoS One* 12, e0182318. doi: 10.1371/journal.pone.0182318
- Skovsted, A. (1935). Chromosome numbers in the malvaceae. *I. J. Genet.* 31, 263–296. doi: 10.1007/BF02982344
- Soltis, D. E., Soltis, P. S., and Tate, J. A. (2004). Advances in the study of polyploidy since plant speciation. *New Phytol.* 161, 173–191. doi: 10.1046/j.1469-8137.2003.00948.x
- Stull, G. W., Pham, K. K., Soltis, P. S., and Soltis, D. E. (2023). Deep reticulation: the long legacy of hybridization in vascular plant evolution. *Plant J.* 114, 743–766. doi: 10.1111/tpj.16142
- Sun, P., Lu, Z., Wang, Z., Wang, S., Zhao, K., Mei, D., et al. (2024). Subgenome-aware analyses reveal the genomic consequences of ancient allopolyploid hybridizations throughout the cotton family. *Proc. Natl. Acad. Sci. U.S.A.* 121, e2313921121. doi: 10.1073/pnas.2313921121
- Tate, J. A., Fuertes Aguilar, J., Wagstaff, S. J., La Duke, J. C., Bodo Slotta, T. A., and Simpson, B. B. (2005). Phylogenetic relationships within the tribe Malveae (Malvaceae, subfamily Malvoideae) as inferred from ITS sequence data. *Am. J. Bot.* 92, 584–602. doi: 10.3732/ajb.92.4.584
- Teng, J., Wang, J., Zhang, L., Wei, C., Shen, S., Xiao, Q., et al. (2022). Paleopolyploidies and genomic fractionation in major Eudicot Clades. *Front. Plant Sci.* 13, 883140. doi: 10.3389/fpls.2022.883140
- Van de Peer, Y., Ashman, T. L., Soltis, P. S., and Soltis, D. E. (2021). Polyploidy: an evolutionary and ecological force in stressful times. *Plant Cell* 33, 11–26. doi: 10.1093/plcell/koaa015
- Van de Peer, Y., Mizrahi, E., and Marchal, K. (2017). The evolutionary significance of polyploidy. *Nat. Rev. Genet.* 18, 411–424. doi: 10.1038/nrg.2017.26
- Vanneste, K., Baele, G., Maere, S., and Van de Peer, Y. (2014). Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous–Paleogene boundary. *Genome Res.* 24, 1334–1347. doi: 10.1101/gr.168997.113
- Vekemans, D., Proost, S., Vanneste, K., Coenen, H., Viaene, T., Ruelens, P., et al. (2012). Gamma paleohexaploidy in the stem lineage of core eudicots: significance for MADS-box gene and species diversification. *Mol. Biol. Evol.* 29, 3793–3806. doi: 10.1093/molbev/mss183
- Vitales, D., García, S., and Dodsworth, S. (2020). Reconstructing phylogenetic relationships based on repeat sequence similarities. *Mol. Phylogenet. Evol.* 147, 106766. doi: 10.1016/j.ympev.2020.106766
- Wan, J. N., Wang, S. W., Leitch, A. R., Leitch, J. J., Jian, J. B., Wu, Z. Y., et al. (2024). The rise of baobab trees in Madagascar. *Nature* 629, 1091–1099. doi: 10.1038/s41586-024-07447-4
- Watanabe, K., Yahara, T., Denda, T., and Kosuge, K. (1999). Chromosomal evolution in the genus *Brachyscome* (Asteraceae, Astereae): statistical tests regarding correlation between changes in karyotype and habit using phylogenetic information. *J. Plant Res.* 112, 145–161. doi: 10.1007/PL00013869
- Wen, D., Yu, Y., Zhu, J., and Nakhleh, L. (2018). Inferring phylogenetic networks using PhyloNet. *Syst. Biol.* 67, 735–740. doi: 10.1093/sysbio/syy015
- Wickens, G. E. (2008). *The Baobabs: Pachycauls of Africa, Madagascar and Australia* (Dordrecht: Springer Netherlands).
- Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., and Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13875–13879. doi: 10.1073/pnas.0811575106
- Yang, W. L., Liu, P. F., Landis, J. B., Zuo, S. Y., Jiang, D. Z., Wen, J., et al. (2025). An integrative framework reveals widespread gene flow during the early diversification of Malvaceae S.L. Available online at: <https://ssrn.com/abstract=5115625> (Accessed September 06, 2025).
- Zhang, R. G., Zhao, H., Conover, J. L., Shang, H. Y., Liu, D. T., Zhou, M. J., et al. (2025). Reticulate allopolyploidy and subsequent dysploidy drive evolution and diversification in the cotton family. *Nat. Commun.* 16, 7480. doi: 10.1038/s41467-025-62644-7
- Zidar, C., and Elisens, W. (2009). Sacred giants: depiction of Bombacoideae on Maya ceramics in Mexico, Guatemala, and Belize. *Econ. Bot.* 63, 119–129. doi: 10.1007/s12231-009-9079-2
- Zizka, A., Carvalho-Sobrinho, J. G., Pennington, R. T., Queiroz, L. P., Alcantara, S., Baum, D. A., et al. (2020). Transitions between biomes are common and directional in Bombacoideae (Malvaceae). *J. Biogeogr.* 47, 1310–1321. doi: 10.1111/jbi.13815