

RESEARCH

Open Access



Plastid phylogenomics illuminates backbone relationships, divergence times and character evolution in Caryophyllaceae

Haiwen Li¹, Shiqiang Song¹, Hans Peter Comes³, Zhaoping Yang^{1*} and Pan Li^{2*}

Abstract

Background The family Caryophyllaceae has traditionally been divided into three subfamilies (i.e. Alsinoideae, Caryophylloideae, Paronychioideae). However, the monophyly of each subfamily has been challenged, while boundaries, phylogenetic relationships, and divergence times of particular tribes and genera still remain to be inferred.

Results Our plastome-derived phylogenies strongly support the monophyly of Caryophyllaceae with a dated crown age of ca. 56.4 Mya (Late Paleocene). While the three traditional subfamilies proved non-monophyletic, relationships among 11 previously recognized tribes were fully resolved. However, in contrast to earlier studies, we identified tribe Eremogoneae (represented by *Eremogone griffithii*) as sister to tribe Sileneae, and the monotypic genus *Thylacospermum* (*T. caespitosum*) as a member of Caryophyllaceae. BBM inferred that the ancestral states of Caryophyllaceae include: presence of stipules; free sepals; absence of petals; five stamens; and utricle or achene fruit.

Conclusions This is the first comprehensive phylogenomic analysis of Caryophyllaceae that strongly supports the subdivision of the family into 11 tribes. While the initial divergence of this family (ca. 56.4 Mya) may be associated with the Paleocene–Eocene Thermal Maximum (PETM; ca. 55.5 Mya), tribes only started to diversify either during the Late Eocene (ca. 35.1–31.6 Mya) or, in particular, during the Late Oligocene to Mid-Miocene (25.5–15.5 Mya). This temporal coincidence suggests that the origin and diversification of many, if not the majority of Caryophyllaceae genera may have been shaped by global cooling and/or aridification events since the Mid-to-Late Miocene as well as the Pliocene–Pleistocene.

Keywords backbone phylogeny, Caryophyllaceae, divergence time, phylogenomics, plastome, tribal classification

*Correspondence:

Zhaoping Yang
yzpzky@163.com
Pan Li
panli@zju.edu.cn

¹College of Life Sciences and Technology, Tarim University, Alar
843300, China

²Key Laboratory of Biodiversity and Environment on the Qinghai-Tibetan
Plateau, Ministry of Education, School of Ecology and Environment,
Xizang University, Lhasa 850000, China

³Department of Environment & Biodiversity, University of Salzburg,
Salzburg, Austria



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Caryophyllaceae Juss. (ca. 100 genera, 3,000 spp.), is the anthocyanidin-producing family of core Caryophyllales, and sister to a clade composed of Achatocarpaceae and Amaranthaceae [1, 2]. Most species of Caryophyllaceae are annual or perennial herbs or subshrubs [3], which are mainly distributed in the temperate and warm temperate regions of the Northern Hemisphere, with centers of diversity in Mediterranean and Irano-Turanian regions [1]; however, a number of species also occur in Africa, South America and Oceania [4]. Many Caryophyllaceae are of medicinal importance [5–7]. For example, in Traditional Chinese Medicine (TCM), several species of this family are used for the treatment of tuberculosis or fever [5]. Others are popular ornamental plants or used in landscaping [6].

Traditionally, three subfamilies have been recognized within Caryophyllaceae, i.e. Alsinoideae, Caryophylloideae (also called Silenoideae) and Paronychioideae, largely based on stipule, flower and fruit as well as embryological characteristics [3]. For example, Paronychioideae have been distinguished from the two other subfamilies by a ‘Solanad-type’ embryo development, often apetalous flowers, and the presence of stipulate leaves, as a likely ancestral trait of the whole family [3, 8]. By contrast, Alsinoideae and Caryophylloideae share a ‘Caryophyllad-type’ embryology as possible synapomorphy but differ in the degree of fusion of the sepals [3]. Moreover, since base chromosome numbers tend to be higher in Alsinoideae ($x = 6–19$) and particularly Caryophylloideae ($x = 10–18$) than in Paronychioideae ($x = 8–9$), it has been suggested that the former two subfamilies may be more closely related or that Caryophylloideae are derived from Alsinoideae [3, 9]. However, early molecular phylogenetic studies already revealed considerable homoplasy in putative diagnostic morphological traits of these three

subfamilies [10, 11], and their non-monophyly became apparent in a molecular systematics study by Harbaugh et al. [12], based on three chloroplast (cp.) DNA regions (*matK*, *trnL-F*, *rps16*) for 126 species. Moreover, these authors were the first to propose abandoning the subfamilies of Caryophyllaceae while recognizing 11 tribes, as also supported by morphological traits [3]. A subsequent phylogeny based on both nuclear ribosomal DNA (ITS) and five cpDNA markers (*matK*, *ndhF*, *trnL-F*, *trnQ-rps16*, *trnS-trnfM*) of 615 species [4] further confirmed the non-monophyly of the three ‘traditional’ subfamilies and broadly agreed with the 11 clades (viz. tribes) proposed by Harbaugh et al. [12].

Nonetheless, although this tribal classification is now widely accepted, boundaries and phylogenetic relationships of some tribes and genera of Caryophyllaceae still remain controversial [4, 12]. For example, in Harbaugh et al. [12], tribe Eremogoneae was identified as sister to a clade composed of Alsineae and Arenarieae, but as sister to Caryophylleae in Greenberg and Donoghue [4]. Besides, the monotypic genus *Thylacospermum* Fenzl (represented by its type species, *T. caespitosum* (Cambess.) Schischk.), a prominent high-alpine cushion plant distributed from western China to the Himalayas and Tianshan Mts [13], was inferred to be either a member of tribe Eremogoneae [12], tribe Sperguleae [4, 14], or as proposed by Prashant [13], tribe Thylacospermeae.

In sum, the monophyly of subfamilies within Caryophyllaceae has been challenged, and controversy still surrounds the boundaries and phylogenetic relationships of particular tribes and genera, such as Caryophylleae, Eremogoneae, *Minuartia* Loebl., etc [2, 4, 12, 14–24]. This taxonomic uncertainty not only hampers a better understanding of the evolutionary history of this family but could also have negative impacts on the conservation and utilization of particular lineages and species [25]. Here in this study, we aimed at generating a robust, time-calibrated backbone phylogeny of Caryophyllaceae, using protein-coding DNA sequence (CDS) data of complete plastomes. The dated and highly supported backbone phylogeny reported here should thus provide a better foundation for downstream macroevolutionary studies on Caryophyllaceae, for instance with focus on the family’s biogeographic history, temporal course of lineage diversification and/or trait evolution (see also [4, 26, 27]).

Materials and methods

Taxon sampling

The plastomes of 23 species of Caryophyllaceae, representing 18 genera, were newly sequenced (Table 1). Next-generation sequencing (NGS) data from 10 species were downloaded from the Sequence Read Archive (SRA) (<https://www.ncbi.nlm.nih.gov/sra>) and assembled into complete plastomes (Table 2). In addition, 46 previously

Table 2 Assembled plastome sequences from NCBI-SRA database for this study

Species name	SRA run number	GenBank accession No.
<i>Arenaria serpyllifolia</i> L.	SRR16841537	OQ476635
<i>Atocion rupestre</i> (L.) Oxelman	ERR5555025	OQ476641
<i>Cherleria arctica</i> (Steven ex Ser.) A.J.Moore & Dillenb.	ERR5554816	OQ476632
<i>Corrigiola litoralis</i> L.	ERR2401628	OQ476640
<i>Geocarpon groenlandicum</i> (Retz.) E.E.Schill.	ERR5529387	OQ476633
<i>Moehringia lateriflora</i> (L.) Fenzl	ERR5555193	OQ476636
<i>Pseudocherleria macrocarpa</i> (Pursh) Dillenb. & Kadereit	ERR5529418	OQ476634
<i>Sabulina rubella</i> (Wahlenb.) Dillenb. & Kadereit	ERR5555283	OQ476637
<i>Saponaria officinalis</i> L.	ERR5529574	OQ476638
<i>Viscaria alpina</i> (L.) G.Don	ERR5554947	OQ476639

Table 1 Voucher and accession numbers of Caryophyllaceae species sampled in this study for plastome sequencing

Species name	Voucher sample no.	Locality	Herbarium	Accession no.
<i>Acanthophyllum pungens</i> (Bunge) Boiss.	LP173592	Bole city, Xinjiang, China	HZU	OP353905
<i>Dianthus armeria</i> L.	King, Robert Merrill 13,935	United States	MO	OP353925
<i>Dianthus nudiflorus</i> Griff.	Merello, Mary Christine 2961	Morocco	MO	OP353926
<i>Drymaria cordata</i> (L.) Willd. ex Roem. & Schult.	Stevens, Warren Douglas 33,226	Nicaragua	MO	OP353906
<i>Eremogone griffithii</i> (Boiss.) Ikonn.	LP173800	Khushikat to Dushanbe, Tajikistan	HZU	OP353907
<i>Gypsophila pacifica</i> Kom.	PNLI20130681	Mishan city, Helongjiang, China	HZU	OP353908
<i>Gypsophila patrinii</i> Ser.	PNLI20130717	Habahe county, Xinjiang, China	HZU	OP353909
<i>Herniaria glabra</i> L.	Yang20170008	Fuhai county, Xinjiang, China	Tarim University	OP353910
<i>Petrorhagia saxifraga</i> (L.) Link	Merello, Mary Christine 2235	Republic of Georgia	MO	OP353913
<i>Pollichia campestris</i> Aiton	Snow, Neil Wilton 7088	South Africa	MO	OP353914
<i>Polycarpaea hassleriana</i> Chodat	Stevens, Warren Douglas 31,338	Paraguay	MO	OP353912
<i>Polycarpon tetraphyllum</i> (L.) L.	Turland, Nicholas John 1871	Greece	MO	OP353915
<i>Pseudostellaria heterantha</i> (Maxim.) Pax	LP161580	Miyun district, Beijing, China	HZU	OP353916
<i>Pseudostellaria sylvatica</i> (Maxim.) Pax	PNLI20141102	Laiyuan county, Hebei, China	HZU	OP353917
<i>Sabulina tenuifolia</i> (L.) Rchb.	Turland, Nicholas John 1754	Greece	MO	OP353911
<i>Sagina japonica</i> (Sw.) Ohwi	Yang2017007	Alar, Xinjiang, China	Tarim University	OP353918
<i>Schiedea hookeri</i> A. Gray	Waimea Arboretum s.n. #26	United States	MO	OP353919
<i>Silene firma</i> Siebold & Zucc.	PNLI20130690	Mudanjiang city, Helongjiang, China	HZU	OP353920
<i>Spergularia marina</i> (L.) Besser	Yang2017009	Hangzhou city, Zhejiang, China	Tarim University	OP353921
<i>Stellaria soongorica</i> Roshev.	PNLI20130731	Fukang city, Xinjiang, China	HZU	OP353922
<i>Stellaria vestita</i> Kurz	LP185503	Hangzhou city, Zhejiang, China	HZU	OP353923
<i>Stellaria yunnanensis</i> Franch.	LP1110123	Kunming city, Yunnan, China	HZU	OP353924
<i>Thylacospermum caespitosum</i> (Cambess.) Schischk.	ZhangDC-001-015	Dingri county, Xizang, China	HZU	Incomplete assembly

published complete plastomes of Caryophyllaceae and 21 of three related families (Achatocarpaceae, Amaranthaceae, and Macarthuriaceae) were downloaded from GenBank (Table 3). Our sampling thus covered 27 genera and all 11 tribes of Caryophyllaceae, as proposed by Harbaugh et al. [12] and Greenberg and Donoghue [4]. For the newly sequenced species (Table 1), leaf material was collected in the field or from herbarium specimens.

DNA extraction, plastome sequencing, assembly and annotation

Total genomic DNAs were extracted from the leaf tissue using a Hi-Fast Plant Genomic Extraction Kit (Gene-Better Biotech Co., Ltd, Beijing, China) and checked for quality and concentration on agarose gel electrophoresis and a NanoDrop2000 spectrometer (Thermo Scientific, United States) to meet the requirements for plastome sequencing. Short-insert (500 bp) paired-end libraries for each sample were constructed using the Genomic DNA Sample Prep Kit (Illumina, San Diego, CA, United States). The genomic DNA of each sample was indexed and these tagged DNAs were then pooled together for sequencing on a HiSeq™ 2500 platform (Illumina) at the Beijing Genomics Institute (BGI, Shenzhen, China).

NOVOPLASTY v. 2.7.2 [28] was used to assemble all 33 plastomes of this study, with the plastome of *Spergula arvensis* L. (GenBank Accession No.: MK330003) serving as a reference [2]. Each assembled plastome was first checked for plastome correctness and locations of genes using GENEIOUS v. 10.0.2 [29], then automatically annotated by PLANN v. 1.1 [30], and finally imported into SEQUIN v. 15.5 (<ftp://ftp.ncbi.nih.gov/sequin/>) to manually add or modify unannotated or incorrectly annotated genes. Graphical maps of each plastome were generated using OrganelleGenomeDRAW v. 1.3.1 [31].

Identification of hypervariable regions in the Caryophyllaceae plastomes

We evaluated levels of nucleotide diversity (Pi) for protein-coding (CDS) sequences vs. noncoding [intergenic spacer (IGS) and intron] sequences across all Caryophyllaceae plastomes of this study to identify hypervariable organellar markers for potential use in future phylogenetic/phylogeographic studies of this family. Values of Pi were calculated by a sliding window approach in DNASP v. 6.0 [32], using a window length of 600 bp and a step size of 200 bp. In addition, we used MAFFT v. 7.2 [33] to compare the sequence order of exons, introns and IGS

regions of the large single-copy (LSC) region, the pair of inverted repeats (IR_A , IR_B), and the small single-copy (SSC) region of each plastome.

Phylogenetic analyses and divergence time estimation

Phylogenetic analyses of Caryophyllaceae and related taxa were performed on shared CDS genes of the 33 newly assembled and 67 downloaded plastomes. Species of Amaranthaceae and Achatocarpaceae were included in the ingroup to test the monophyly of Caryophyllaceae and to incorporate additional fossil calibration points, while *Macarthuria keigheryi* Lepschi (Macarthuraceae) served as outgroup [2] (Table 3). PHYLOSUITE v. 1.2.2 [34] was used to derive CDS genes for each plastome. The 77 CDS sequences shared between the 100 plastomes were concatenated and aligned with the L-INS-i algorithm implemented in MAFFT v. 7.2. The resulting matrix was subjected to phylogenetic analyses, using both Maximum Likelihood (ML) in RAXML-HPC2 v. 7.6.3 [35] and Bayesian Inference (BI) in MRBAYES v. 3.2.2 [36], as executed via XSEDE tools on the CIPRES Science Gateway website [37]. Preliminary evaluation of the dataset revealed no significant heterogeneity in evolutionary rates or nucleotide substitution patterns among CDS genes. We also tested and compared partitioning strategies based on genes and codon positions, and the results showed that the choice of partitioning schemes had minimal impact on the inferred topology and key branch support values (e.g., BS, PP) of the backbone relationships. Consequently, instead of implementing partition models by gene or codon position, we employed the GTR + G+I model of nucleotide substitution for the entire dataset, as inferred by the Akaike Information Criterion (AIC) in jModelTest v. 2.1.7 [38]. The ML analysis used 1,000 ‘fast bootstrap’ replicates. The Markov chain Monte Carlo (MCMC) simulations of the BI analysis were run with two independent chains set for 1,000,000 generations, and the first 25% of trees from all runs were discarded as burn-in. Finally, we employed BEAST v. 2.5.2 [39] with the GTR + G model to generate a dated phylogeny from the full CDS dataset of the 100 plastomes under the assumption of a relaxed clock. The tree speciation prior was set to follow a Yule process, and three fossil dates were used for calibration (see below). Five independent MCMC chains were run for 80,000,000 generations, while sampling every 8,000 generations. The initial 10% of trees from all runs were discarded as burn-in. The five runs were combined using LOGCOMBINER v. 2.5.2 [39], and checked for convergence [effective sample size (ESS) > 200] using TRACER v. 1.5 (<https://tree.bio.ed.ac.uk/software/tracer>). The resulting maximum clade credibility (MCC) chronogram with mean node ages and 95% highest posterior density (HPD) intervals was annotated and visualized using TREEANNOTATOR v. 2.5

[39] and FIGTREE v. 1.4.0 (<http://tree.bio.ed.ac.uk/software/figtree>), respectively.

Despite the generally sparse fossil record of Caryophyllaceae [40–43], we employed three calibration points (C1–C3) in our BEAST analysis based on plastome-derived CDS genes. First, referring to studies of Kadereit et al. [44–46], Li et al. [47] and Yao et al. [2], the crown node of Amaranthaceae (C1) was assigned a log-normal prior (offset = 66.0 Mya; mean = 1.0, SD = 0.5). This minimum age is based on the late-Cretaceous pollen fossil *Polyporina cribaria* Srivastava [48], a widely accepted early representative of the family, and the prior allows for a plausible older age while weighting values near the fossil constraint. In addition, the crown node of Caryophyllaceae (C2) was given a normal prior centered at 57.0 Mya (mean = 1.0, SD = 1.0), with 57.0 Mya set as the hard minimum. This point follows the integrated dating result of Yao et al. [2], which is consistent with the earliest known macrofossils of the family. Finally, the crown of tribe Sileneae (C3) was calibrated with a log-normal prior (offset = 20.8 Mya; mean = 0.5, SD = 1.0) based on the estimate of Mahmoudi Shamsabad et al. [22]; the offset corresponds to a conservative minimum age congruent with the appearance of Silene-type pollen in the Miocene. Together, these points incorporate key fossil evidence and published estimates while using explicit prior distributions to accommodate age uncertainty, thereby improving the robustness of divergence-time estimates given the limited direct fossil record for Caryophyllaceae.

Morphological character evolution

For ancestral state reconstructions within Caryophyllaceae, we selected five morphological characters that have generally been considered distinguishing traits among the three traditional subfamilies [3]: stipule presence/absence, calyx tube presence/absence, petal presence/absence, stamen number (5 vs. 10), and fruit type (capsule vs. achene or utricle). Following the approach of Greenberg and Donoghue [4], each of these characters was treated as a discrete, binary trait. A single, explicitly defined coding scheme was applied per character. Character states were defined based on standard taxonomic descriptions and were assigned through direct examination of herbarium specimens and review of authoritative monographs [3, 4, 8]. The ancestral states of these five characters were then reconstructed using BBM (Bayesian Binary MCMC) analysis in RASP v.4.2 [49], based on the BEAST-derived MCC tree topology of Caryophyllaceae.

Results

Characteristics of the newly sequenced Caryophyllaceae plastomes

We newly sequenced the plastomes of 23 species of Caryophyllaceae and provide high-quality plastome

Table 3 Plastome sequences obtained from NCBI for this study

Family	Species	Accession no.
Achatocarpaceae	<i>Achatocarpus nigricans</i> Triana	MK397908.1
Achatocarpaceae	<i>Achatocarpus pubescens</i> C.H.Wright	MK397909.1
Achatocarpaceae	<i>Phaulothamnus spinescens</i> A.Gray	MH286322.1
Amaranthaceae	<i>Achyranthes aspera</i> L.	NC_050063.1
Amaranthaceae	<i>Aerva javanica</i> (Burm.f.) Juss. ex Schult.	MK410028.1
Amaranthaceae	<i>Alternanthera philoxeroides</i> (Mart.) Griseb.	NC_042798.1
Amaranthaceae	<i>Amaranthus hybridus</i> E.H.L.Krause	NC_053787.1
Amaranthaceae	<i>Celosia cristata</i> L.	NC_045887.1
Amaranthaceae	<i>Chenopodium acuminatum</i> Willd.	NC_054154.1
Amaranthaceae	<i>Chenopodium album</i> L.	MW417304.1
Amaranthaceae	<i>Cyathula capitata</i> Wall. ex Moq.	NC_041262.1
Amaranthaceae	<i>Deeringia amaranthoides</i> (Lam.) Merr.	NC_041267.1
Amaranthaceae	<i>Froelichia latifolia</i> R.A.McCauley	MH286309.1
Amaranthaceae	<i>Gomphrena cunninghamii</i> (Moq.) Druce	MK410014.1
Amaranthaceae	<i>Omegandra kanisii</i> G.J.Leach & C.C.Towns.	MK410034.1
Amaranthaceae	<i>Ourea glabrata</i> (Hook.f.) Kuntze	MK410031.1
Amaranthaceae	<i>Oxybasis glauca</i> (L.) S.Fuentes, Uotila & Borsch	NC_047226.1
Amaranthaceae	<i>Paraerva microphylla</i> (Moq.) T.Hammer	MK410032.1
Amaranthaceae	<i>Ptilotus polystachyus</i> (Gaudich.) F.Muell.	NC_046575.1
Amaranthaceae	<i>Suaeda salsa</i> (L.) Pall.	NC_045302.1
Caryophyllaceae	<i>Agrostemma githago</i> L.	NC_023357.1
Caryophyllaceae	<i>Cerastium arvense</i> L.	MH627219.1
Caryophyllaceae	<i>Colobanthus acicularis</i> Hook.f.	NC_053724.1
Caryophyllaceae	<i>Colobanthus affinis</i> (Hook.) Hook.f.	NC_053722.1
Caryophyllaceae	<i>Colobanthus apetalus</i> (Labill.) Druce	NC_036424.1
Caryophyllaceae	<i>Colobanthus lycopodioides</i> Griseb.	NC_053721.1
Caryophyllaceae	<i>Colobanthus nivicola</i> M.Gray	NC_053720.1
Caryophyllaceae	<i>Colobanthus pulvinatus</i> F.Muell.	NC_053719.1
Caryophyllaceae	<i>Colobanthus quitensis</i> (Kunth) Bartl.	NC_028080.1
Caryophyllaceae	<i>Colobanthus subulatus</i> (d'Urv.) Hook.f.	NC_053723.1
Caryophyllaceae	<i>Dianthus caryophyllus</i> L.	NC_039650.1
Caryophyllaceae	<i>Dianthus chinensis</i> L.	NC_053731.1
Caryophyllaceae	<i>Dianthus gratianopolitanus</i> Vill.	LN877394.1
Caryophyllaceae	<i>Dianthus longicalyx</i> Miq.	NC_050834.1
Caryophyllaceae	<i>Dianthus moravicus</i> Kovanda	LN877396.1
Caryophyllaceae	<i>Gymnocarpus przewalskii</i> Bunge ex Maxim.	NC_036812.1
Caryophyllaceae	<i>Gypsophila vaccaria</i> (L.) Sm.	NC_040936.1
Caryophyllaceae	<i>Lychnis wilfordii</i> (Regel) Maxim.	NC_035225.1
Caryophyllaceae	<i>Psammosilene tunicoides</i> W.C.Wu & C.Y.Wu	NC_045947.1
Caryophyllaceae	<i>Pseudocerasium stellarioides</i> X.H.Guo & X.P.Zhang	MT507771.1
Caryophyllaceae	<i>Pseudostellaria heterophylla</i> (Miq.) Pax	NC_044183.1
Caryophyllaceae	<i>Pseudostellaria longipedicellata</i> S.Lee, K.I.Heo & S.C.Kim	NC_039454.1
Caryophyllaceae	<i>Pseudostellaria okamotoi</i> Ohwi	NC_039974.2
Caryophyllaceae	<i>Pseudostellaria palibiniana</i> (Takeda) Ohwi	NC_041166.1
Caryophyllaceae	<i>Pseudostellaria setulosa</i> Ohwi	NC_041462.1
Caryophyllaceae	<i>Silene aprica</i> Turcz. ex Fisch. & C.A.Mey.	NC_040934.1
Caryophyllaceae	<i>Silene cambessedesii</i> Boiss. & Reut.	MN365977.1
Caryophyllaceae	<i>Silene capitata</i> Kom.	NC_035226.1
Caryophyllaceae	<i>Silene chalcidonica</i> (L.) E.H.L.Krause	NC_023359.1
Caryophyllaceae	<i>Silene conica</i> L.	NC_016729.1
Caryophyllaceae	<i>Silene conoidea</i> L.	NC_023358.1
Caryophyllaceae	<i>Silene jensseensis</i> Willd.	MN723869.1
Caryophyllaceae	<i>Silene kiusiana</i> (Makino) H.Ohashi & H.Nakai	NC_048886.1

Table 3 (continued)

Family	Species	Accession no.
Caryophyllaceae	<i>Silene latifolia</i> Poir.	NC_016730.1
Caryophyllaceae	<i>Silene latifolia</i> subsp. <i>alba</i> (Mill.) Greuter	MN244687.1
Caryophyllaceae	<i>Silene littorea</i> Brot.	MN365987.1
Caryophyllaceae	<i>Silene littorea</i> subsp. <i>adscendens</i> (Lag.) Rivas Goday	MN365972.1
Caryophyllaceae	<i>Silene noctiflora</i> L.	NC_016728.1
Caryophyllaceae	<i>Silene paradoxa</i> L.	NC_023360.1
Caryophyllaceae	<i>Silene psammitis</i> Link ex Spreng.	MN365988.1
Caryophyllaceae	<i>Silene stockenii</i> Chater	MN365992.1
Caryophyllaceae	<i>Silene uniflora</i> Roth	KY562597.1
Caryophyllaceae	<i>Silene vulgaris</i> (Moench) Garcke	NC_016727.1
Caryophyllaceae	<i>Silene vulgaris</i> subsp. <i>vulgaris</i> (Moench) Garcke	MK473866.1
Caryophyllaceae	<i>Spergula arvensis</i> L.	NC_041240.1
Caryophyllaceae	<i>Stellaria dichotoma</i> var. <i>lanceolata</i> Bunge	MW442089.1
Marcathaceae	<i>Macarthuria keigheryi</i> Lepschi	MK397926.1

Table 4 A comparison of the 22 newly sequenced and assembled plastomes of Caryophyllaceae representing 18 genera and 22 species. Note that species are ordered according to the percentage of total plastome length covered by protein-coding (CDS) genes (last column)

Species	Plastome size (bp)	LSC (bp)	IR (bp)	SSC (bp)	Predicted functional genes	Protein-coding (CDS) genes (% of total length)
<i>Acanthophyllum pungens</i> (Bunge) Boiss.	150,491	83,231	24,882	17,366	128	84 (52.12)
<i>Dianthus armeria</i> L.	150,120	83,021	24,886	17,327	128	84 (53.30)
<i>Dianthus nudiflorus</i> Griff.	149,676	82,868	24,764	17,280	127	82 (54.31)
<i>Drymaria cordata</i> (L.) Willd. ex Roem. & Schult.	151,288	82,344	25,981	16,997	127	84 (55.05)
<i>Eremogone griffithii</i> (Boiss.) Ikonn.	135,008	74,864	22,002	16,140	129	85 (56.54)
<i>Gypsophila pacifica</i> Kom.	152,778	83,741	25,775	17,487	127	83 (52.05)
<i>Gypsophila patrinii</i> Ser.	152,681	83,647	25,845	17,344	127	83 (52.22)
<i>Herniaria glabra</i> L.	153,102	84,295	26,043	16,687	127	83 (52.83)
<i>Petrorhagia saxifraga</i> (L.) Link	148,413	81,519	24,908	17,084	127	82 (54.29)
<i>Pollichia campestris</i> Aiton	153,615	83,767	26,248	17,331	127	82 (53.73)
<i>Polycarpaea hassleriana</i> Chodat	151,475	83,205	25,733	16,804	127	82 (53.82)
<i>Polycarpon tetraphyllum</i> (L.) L.	151,811	81,844	26,342	17,283	128	83 (54.69)
<i>Pseudostellaria heterantha</i> (Maxim.) Pax	149,672	81,266	25,681	17,044	127	83 (53.85)
<i>Pseudostellaria sylvatica</i> (Maxim.) Pax	149,581	81,089	25,774	16,944	127	83 (53.90)
<i>Sabulina tenuifolia</i> (L.) Rchb.	152,715	83,567	26,174	16,879	127	82 (51.23)
<i>Sagina japonica</i> (Sw.) Ohwi	151,449	83,633	25,311	17,194	128	84 (53.17)
<i>Schiedea hookeri</i> A. Gray	146,389	84,758	22,464	16,703	126	83 (55.52)
<i>Silene firma</i> Siebold & Zucc.	150,242	82,230	25,344	17,324	127	83 (52.73)
<i>Spergularia marina</i> (L.) Besser	152,395	83,285	25,890	17,330	127	82 (54.13)
<i>Stellaria soongorica</i> Roshev.	148,342	80,625	25,494	16,729	127	83 (51.35)
<i>Stellaria vestita</i> Kurz	148,342	80,622	25,578	16,585	127	83 (53.57)
<i>Stellaria yunnanensis</i> Franch.	148,539	80,682	25,584	16,691	127	82 (52.43)

assemblies for 22 of them (Table 4; Fig. 1). However, for *Thylacospermum caespitosum*, we could only assemble the all shared CDS genes due to otherwise poor plastome-sequencing quality. The 22 newly sequenced and fully assembled plastomes possessed the typical quadripartite structure of most land plants (Table 4; Fig. 1), consisting of a pair of IR regions (IR_A, IR_B), separated by the LSC and SSC regions. Yet, these plastomes differed significantly in overall size (Table 4), ranging

from 135,008 bp [*Eremogone griffithii* (Boiss.) Ikonn.] to 153,615 bp (*Pollichia campestris* Aiton). Likewise, each of the four regions (IR_A, IR_B, LSC, SSC) varied in size, with the LSC ranging from 74,864 bp (*E. griffithii*) to 84,758 bp (*Schiedea hookeri* A. Gray), the SSC from 16,140 bp (*E. griffithii*) to 17,487 bp (*Gypsophila pacifica* Kom.), and the IRs from 22,002 bp (*E. griffithii*) to 26,342 bp [*Polycarpon tetraphyllum* (L.) L.] (Table 4).

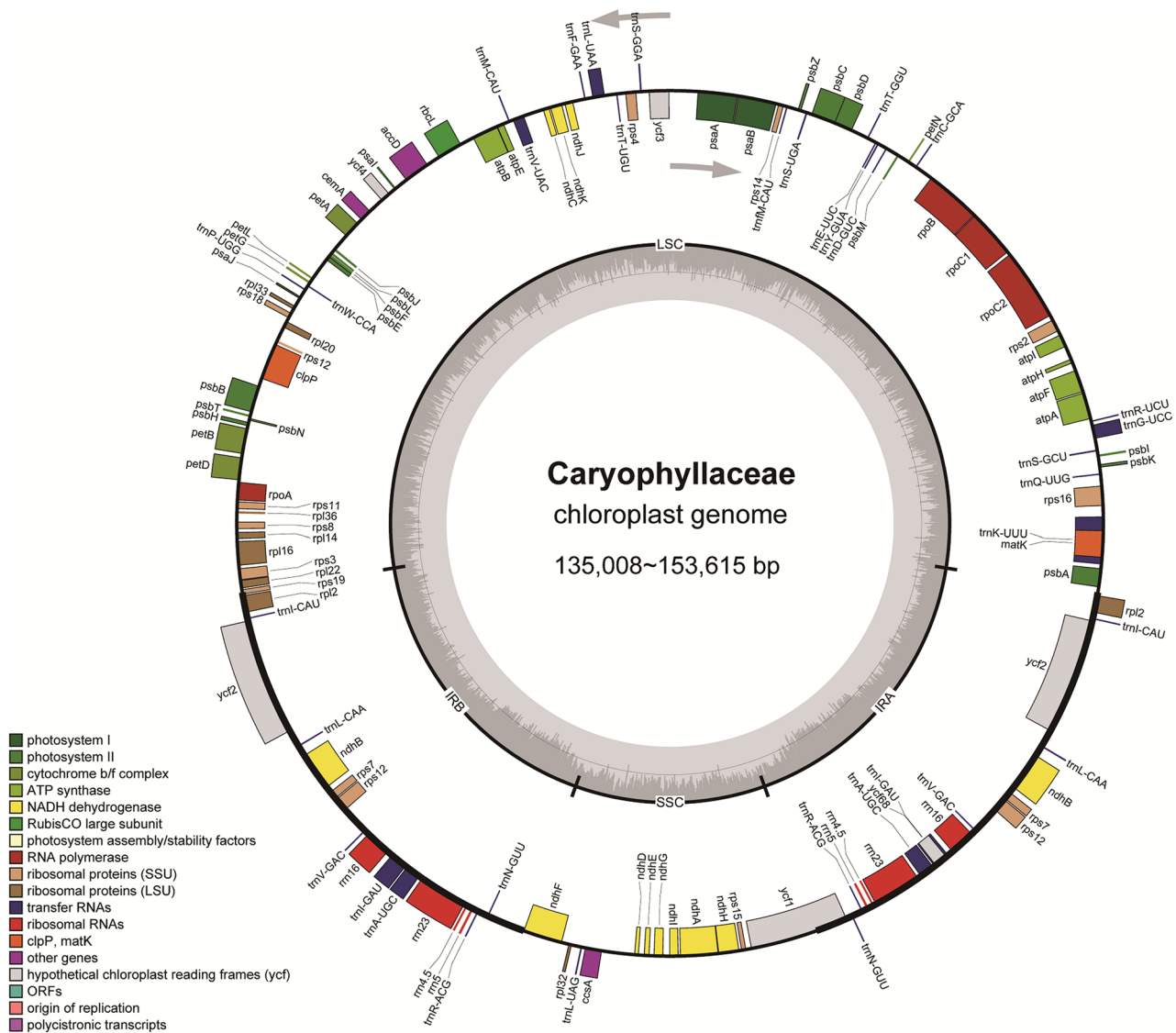


Fig. 1 Circular chloroplast genome (plastome) map of Caryophyllaceae based on the plastomes of 22 species (size range: 135,008–153,615 bp). Genes drawn outside the circle are transcribed counterclockwise, and those inside are transcribed clockwise. Different colored bars represent genes with different functions. The dashed dark gray area in the inner circle indicates the GC content of the plastome, and the light gray area shows the AT content. Abbreviations: LSC, large single-copy region; SSC, small single-copy region; IR_A and IR_B, inverted repeat regions

Among the 22 complete plastomes, the total number of predicted functional genes ranged from 126 to 129, including 82 [*Sabulina tenuifolia* (L.) Rchb.] to 85 [*Eremogone griffithii*] CDS genes, which accounted for 51.23% to 56.54% of the total plastome length (Table 4). When compared to the other species, the strikingly small plastome of *E. griffithii* (135,008 bp; see above) was mainly due the shortening of genes and IGS regions (e.g. *ycf1*, *ycf2*, *ycf3*, *rpl16*, *rps7-trnV*, *psbE-petL*, *petN-psbM*) or even the entire absence of particular IGS sequences (e.g. *accD-psaI*, *atpB-rbcL*, *trnN-ndhF*, *psaA-ycf3*).

Identification of hypervariable regions

Across the full 79-plastome dataset of Caryophyllaceae, we identified 13 regions with considerably high nucleotide diversity ($P_i > 0.1$) (Fig. 2). Most of these hypervariable regions (11 out of 13) were distributed in the LSC region (i.e. *trnK-rps16*, *psbC-trnS*, *trnS-rps4*, *trnT-trnL*, *trnF-ndhI*, *psaI-ycf4*, *petA-psbJ*, *psaJ-rpl33*, *rpl36-infA*, *infA*, *rpl16-rps3*), yet only two in the SSC region (i.e., *ccsA-ndhD*, *ndhG-ndhI*). Notably, the *infA* gene, coding for a translation initiation factor, had the highest diversity ($P_i = 0.175$) (see Fig. 2).

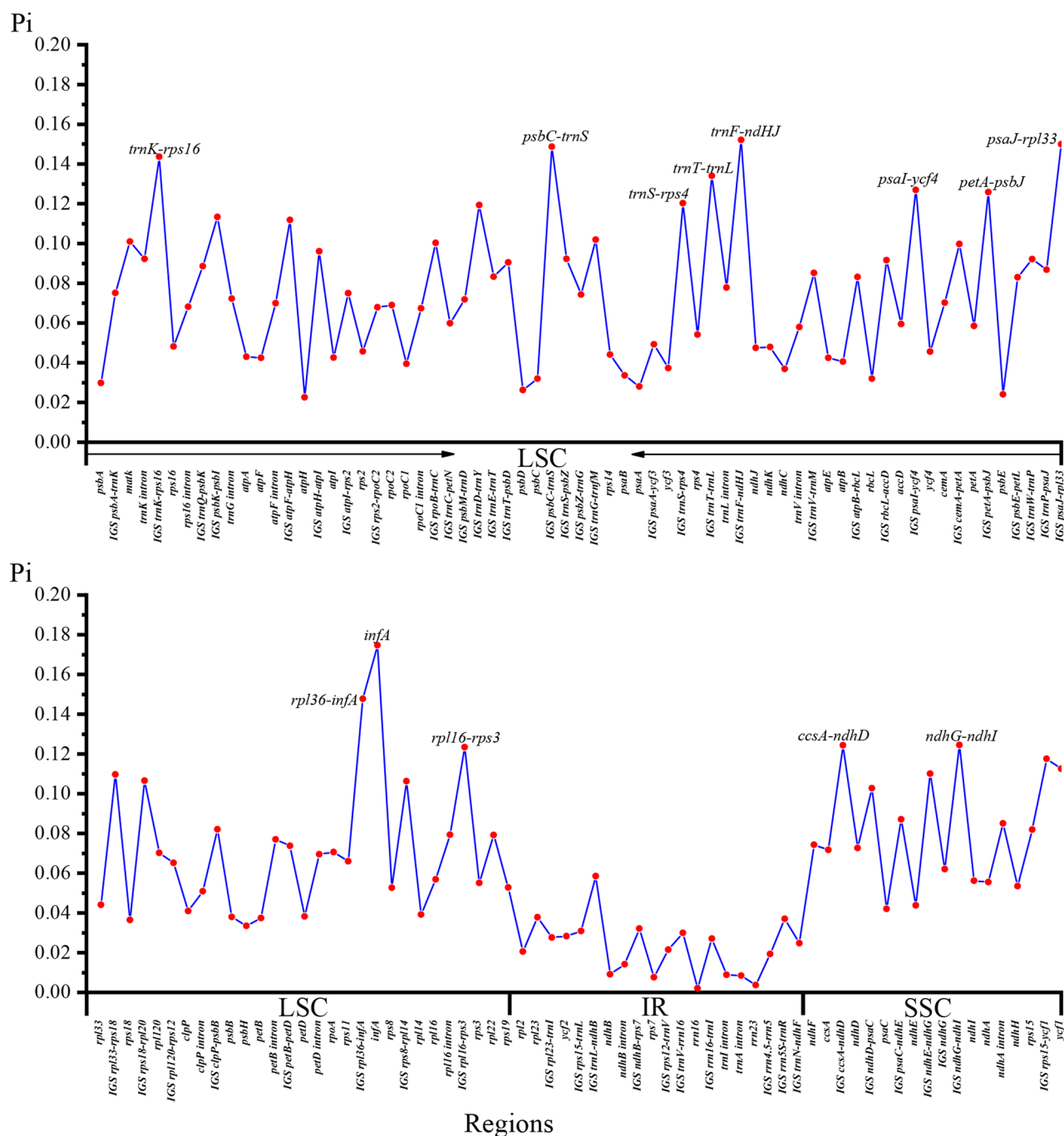


Fig. 2 Comparison of nucleotide diversity (π) of plastomes in Caryophyllaceae. π : nucleotide diversity

Phylogenetic analyses

We obtained a well-supported plastome phylogeny of Caryophyllaceae (79 spp., 38 genera, 11 tribes) and related taxa (Amaranthaceae: 17 spp.; Achatocarpaceae: 3 spp.; Macarthuraceae: 1 sp.) based on 77 shared CDS genes (Fig. 3). As the tree topologies of the ML and BI analyses were identical, we annotated the tree with both ML-bootstrap (BS) and BI-posterior probability (PP) values (Fig. 3). This phylogeny strongly supported the

monophyly of Caryophyllaceae (BS = 100%; PP = 1) as sister to a clade consisting of Amaranthaceae + Achatocarpaceae (both BS = 100; PP = 1). The phylogeny clarified relationships among the three traditional subfamilies: Paronychioideae (P; including tribes Corrigoieae, Paronychieae, Polycarpaeae and Sperguleae) formed an early-diverging grade, whereas Alsinoideae (A; including tribes Alsineae, Arenarieae, Sagineae and Scleranthaeae) and Caryophylloideae (C; including tribes Caryophylleae and

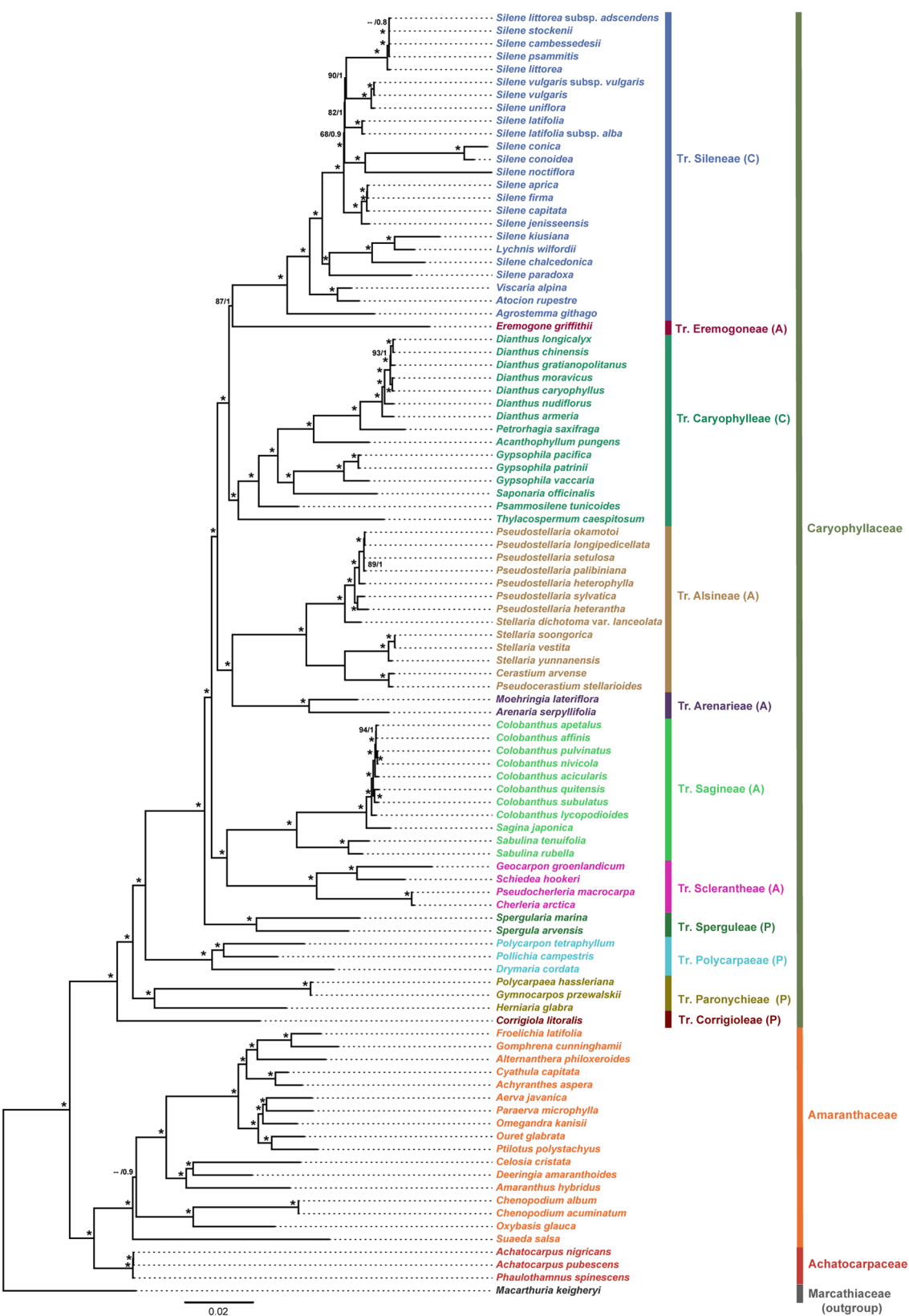


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Phylogenetic tree reconstruction of Caryophyllaceae (79 spp., 38 genera, 11 tribes) and related taxa (Amaranthaceae: 17 spp.; Achatocarpaceae: 3 spp.; Marcathuriaceae: 1 sp.) based on 77 protein-coding (CDS) genes inferred from Maximum Likelihood (ML) and Bayesian Inference (BI) analyses. Numbers at each node represent ML bootstrap support (BS) values and BI posterior probability (PP) values, with asterisks indicating BS = 100% and PP = 1.0. For Caryophyllaceae, letters in parentheses show the assignment of each tribe (Tr.) to a traditionally recognized subfamily (A: Alsinoideae; C: Caryophylloideae; P: Paronychioideae)

Sileneae) together constituted a highly supported monophyletic group (BS = 100%; PP = 1). However, within this clade, Alsinoideae (excluding Eremogoneae) was paraphyletic with respect to Caryophylloideae, which itself included the Alsinoideae tribe Eremogoneae (represented by *Eremogone griffithii*) as sister to Sileneae (BS = 87%; PP = 1). Nonetheless, all tribes of Caryophyllaceae represented by multiple species (i.e. nine out of 11; *sensu* [4, 12]), received strong support (BS > 85%; PP = 1). Notably too, the monotypic genus *Thylacospermum* (*T. caespitosum*) was strongly supported as earliest-diverging lineage of the Caryophylloideae tribe Caryophylleae (BS = 100%; PP = 1) (Fig. 3).

Molecular divergence time estimation

Figure 4 shows the BEAST-derived chronogram of Caryophyllaceae (plus related taxa), and Table 5 summarizes the mean ages (and HPD intervals) for relevant nodes (1–12 in Fig. 4). Accordingly, we dated the stem age of Caryophyllaceae to the Late Cretaceous, ca. 78.7 Mya (95% HPD: 72.0–85.6 Mya; node 1), and their crown age to the Late Paleocene, ca. 56.4 (54.4–58.2) Mya (node 2). While the mean crown ages of particular tribes fell into the Late Eocene (Paronychieae: ca. 35.1 Mya, node 3) or Early to Late Oligocene (Caryophylleae: ca. 31.6 Mya, node 5; Polycarpaeae: ca. 25.5 Mya, node 6), the majority of tribes started to diversify during the Early Miocene (Sileneae: ca. 21.8 Mya, node 7; Scleranthaeae: ca. 21.6 Mya, node 8; Alsinoeae: ca. 20.7 Mya, node 9; Saginoeae: ca. 20.6 Mya, node 10; Sparguleae: ca. 19.5 Mya, node 11), and only rarely thereafter, during the Mid-Miocene (Arenarieae: ca. 15.5 Mya; node 12). Notably, *Thylacospermum caespitosum* diverged from the rest of Caryophylleae already during the Early Oligocene (ca. 31.6 Mya, node 5), nearly simultaneous with the split between Eremogoneae (*Eremogone griffithii*) and Sileneae (ca. 33.7 Mya, node 4).

Ancestral state reconstructions

According to the RASP (BBM) analysis (Fig. 5), we identified the following five ancestral states in Caryophyllaceae: presence of stipules; free sepals; absence of petals; five stamens; and utricle or achene fruit. All five ancestral states were thus present in the most recent common ancestor (MRCA) of the family (see crown node in Fig. 5), Fig. 5 Ancestral state reconstructions of five morphological characters along the BEAST-derived tree topology of Caryophyllaceae (see Fig. 4) using BBM (Bayesian Binary

MCMC) analysis in RASP. A: stipules; B: calyx tube; C: petals; D: stamen number; E: fruit. Numbers at selected nodes (1–6) represent the relative probability (*p*) of occurrence of different ancestral states at these nodes. Letters in parentheses show the assignment of each tribe (Tr.) to a traditionally recognized subfamily (A: Alsinoideae; C: Caryophylloideae; P: Paronychioideae).

and all of them are retained in the three earliest-diverging tribes of Paronychioideae (i.e. Corrigioleae, Paronychieae, and Polycarpaeae). For the other eight tribes of the family (node 1 in Fig. 5), BBM identified the presence of petals, 10 stamens, and capsules as ancestral states (each with relative probability *p* = 0.93). While stipules are still present in four tribes of Paronychioideae (i.e. Corrigioleae, Paronychieae, Polycarpaeae, and Sparguleae), they were probably lost in the MRCA of the seven other tribes (node 2, *p* = 0.89). The node at which a shift from free sepals to a tubular calyx occurred is less clear, as Eremogoneae (sister to Sileneae) have free sepals while Caryophylleae and Sileneae have a tubular calyx. It is possible that this shift occurred twice, once along the branch leading to Caryophylleae (node 5), and once along the branch leading to Sileneae (node 6). However, it is also possible that this character was gained only once, along the branch leading to the Caryophylleae–Eremogoneae–Sileneae clade (node 3), and then was lost in Eremogoneae (see Fig. 5).

Discussion

Plastome-derived backbone phylogeny of Caryophyllaceae

In this study, we generated ML, BI and time-calibrated (BEAST) phylogenies of Caryophyllaceae based on 77 shared CDS genes of 100 plastid genomes, representing all 11 tribes of Caryophyllaceae as well as closely related families (Achatocarpaceae, Amaranthaceae, Marcathuriaceae; Figs. 3 and 4). Our ML/BI tree topologies (Fig. 3) strongly support the monophyly of Caryophyllaceae as well as their sister relationship to a clade comprising Amaranthaceae + Achatocarpaceae. In support of earlier studies [4, 12, 50], our phylogenies (Fig. 4) further indicate that none of the three traditional subfamilies of Caryophyllaceae is monophyletic. While Paronychioideae form an early-diverging grade, Alsinoideae, excluding tribe Eremogoneae, forms a paraphyletic group relative to Caryophylloideae, which includes tribe Eremogoneae as sister to Sileneae. Nonetheless, all tribes of Caryophyllaceae represented in this study by multiple species (nine out of 11, except Corrigioleae and Eremogoneae) were

recovered as strongly supported monophyletic groups (Fig. 3). Hence, this tribal classification scheme appears to be consistent with previous phylogenetic studies [4, 12, 50].

Nevertheless, in those previous studies, phylogenetic relationships among a few tribes were poorly supported [4, 12], with a major conflict involving the phylogenetic position of tribe Eremogoneae. For example, Harbaugh et al. [12] inferred Eremogoneae as sister to a clade comprising Alsineae + Arenarieae, and which together were identified as sister group to a clade including Caryophylleae + Sileneae. In contrast, Greenberg and Donoghue [4] proposed Eremogoneae as sister to Caryophylleae, together forming a clade with Sileneae. Our plastome-derived ML/BI phylogenies (Fig. 3) conflict with both of the aforementioned topologies. Instead, they robustly support a sister-group relationship between tribes Eremogoneae and Sileneae (BS = 87%; PP = 1.0). This result aligns with the phylogenetic relationships recovered in the transcriptome-based study by Feng et al. [50]. These conflicting phylogenetic relationships may be attributed to differences in datasets, as limited marker combinations provide insufficient phylogenetic information to resolve phylogenetic relationships.

Although previous studies [4, 11, 12, 50] also provided a deeper understanding of the phylogenetic relationships of most genera of Caryophyllaceae, some of these relationships remain controversial. As a case in point, Harbaugh et al. [12] proposed that *Thylacospermum* (together with *Eremogone*) should constitute their new tribe Eremogoneae, while Greenberg and Donoghue [4] found this monotypic genus (i.e. *Thylacospermum caespitosum*) nested within Sperguleae. Besides, in a recent study focusing on cushion-plant radiations in the Qinghai-Tibetan Plateau [51], *Thylacospermum* was found to be sister to *Arenaria* subg. *Dolophragma*, together forming a heterogeneous clade with species of Arenarieae, Alsineae, Sagineae and Scleranthaeae, but with generally weak supports. However, our results strongly support *Thylacospermum caespitosum* as earliest-diverging lineage of Caryophylleae (Fig. 3). Overall, these conflicting results imply that further research is needed to unravel the phylogenetic position and evolutionary history of this Asian high-alpine species.

While plastid genomes provide a robust signal, they may be insufficient to fully resolve deep backbone relationships in groups with complex histories involving hybridization or incomplete lineage sorting. The plastid phylogenomic relationships recovered in our study show key disagreements with those of the transcriptome-based analysis by Feng et al. [50]. Specifically, our data support a sister-group relationship between tribes Alsineae and Arenarieae, whereas their results suggest a paraphyletic relationship. Similarly, our analysis resolves tribes

Eremogoneae, Sileneae, and Caryophylleae as monophyletic, while their topology also indicates paraphyly for these tribes. These results indicate potential cytonuclear discordance within certain lineages of Caryophyllaceae, suggesting they may be challenging regions of the tree requiring further investigation with genomic-scale nuclear data.

Divergence time estimation of Caryophyllaceae

Although the extant species of Caryophyllaceae are abundant and diverse, most of them are tender herbs, which leads to the rarity of fossils in this family [40, 41] and a lack of study on molecular dating. The oldest known fossils of Caryophyllaceae are pollen grains from the Late Cretaceous (Campanian, ca. 73 Mya), found in Australia and New Zealand [52, 53]. However, their high morphological similarity to many other taxonomic groups (e.g. Myrtaceae and Casuarinaceae) might render their assignment to this family questionable [54], while the oldest fossilized seeds of Caryophyllaceae (*Hantsia pulchra*) are dated to the Eocene [55]. In this study, we employed a Late Cretaceous pollen record of Amaranthaceae (*Polyporina cribaria*; [48]) plus previous age estimates of Caryophyllaceae [2] and Sileneae [22] to calibrate our BEAST-derived MCC tree (see nodes labelled 'C1–C3' in Fig. 4). According to this chronogram, Caryophyllaceae diverged from the sister clade (Achatocarpaceae + Amaranthaceae) during the Late Cretaceous, ca. 78.7 (72.0–85.6) Mya. The 95% HPD interval around this estimate thus encompasses the oldest fossil pollen record of Caryophyllaceae (ca. 73 Mya; see above). In turn, we dated the crown age of Caryophyllaceae to the Late Paleocene (ca. 56.4 Mya). Notably, the fossil inflorescence *Caryophylloflora paleogenica* G.J.Jord. & Macphail (Middle–Late Eocene; 48.6–33.9 Mya) [42] has been hypothesized to be sister to Alsinoideae and/or Caryophylloideae [42, 56]. Although this placement requires further confirmation, our estimated crown age for the Alsinoideae–Caryophylloideae clade (ca. 40.7 Mya; see Fig. 4) is broadly consistent with the age of this fossil.

Although there are previous studies on divergence time estimates in Caryophyllaceae, these most relate to specific tribes and genera. For example, based on *matK* sequence data of Caryophylloideae and Alsinoideae, Frajman et al. [56] estimated the crown ages of Arenarieae (ca. 13.85 Mya), a clade composed of *Cerastium* and *Stellaria* in Alsineae (ca. 10 Mya), a clade composed of *Dianthus*, *Gypsophila* and *Saponaria* in Caryophylleae (ca. 20.18 Mya), and Sperguleae (ca. 23.34 Mya). All these estimates are broadly consistent with those of the present study (ca. 15.5 Mya, 12.2 Mya, 22.4 Mya, 19.5 Mya, respectively). Similarly, using *rps16* sequence data, Mahmoudi Shamsabad et al. [22] studied a clade composed of *Dianthus*, *Gypsophila* and *Saponaria* in Caryophylleae,

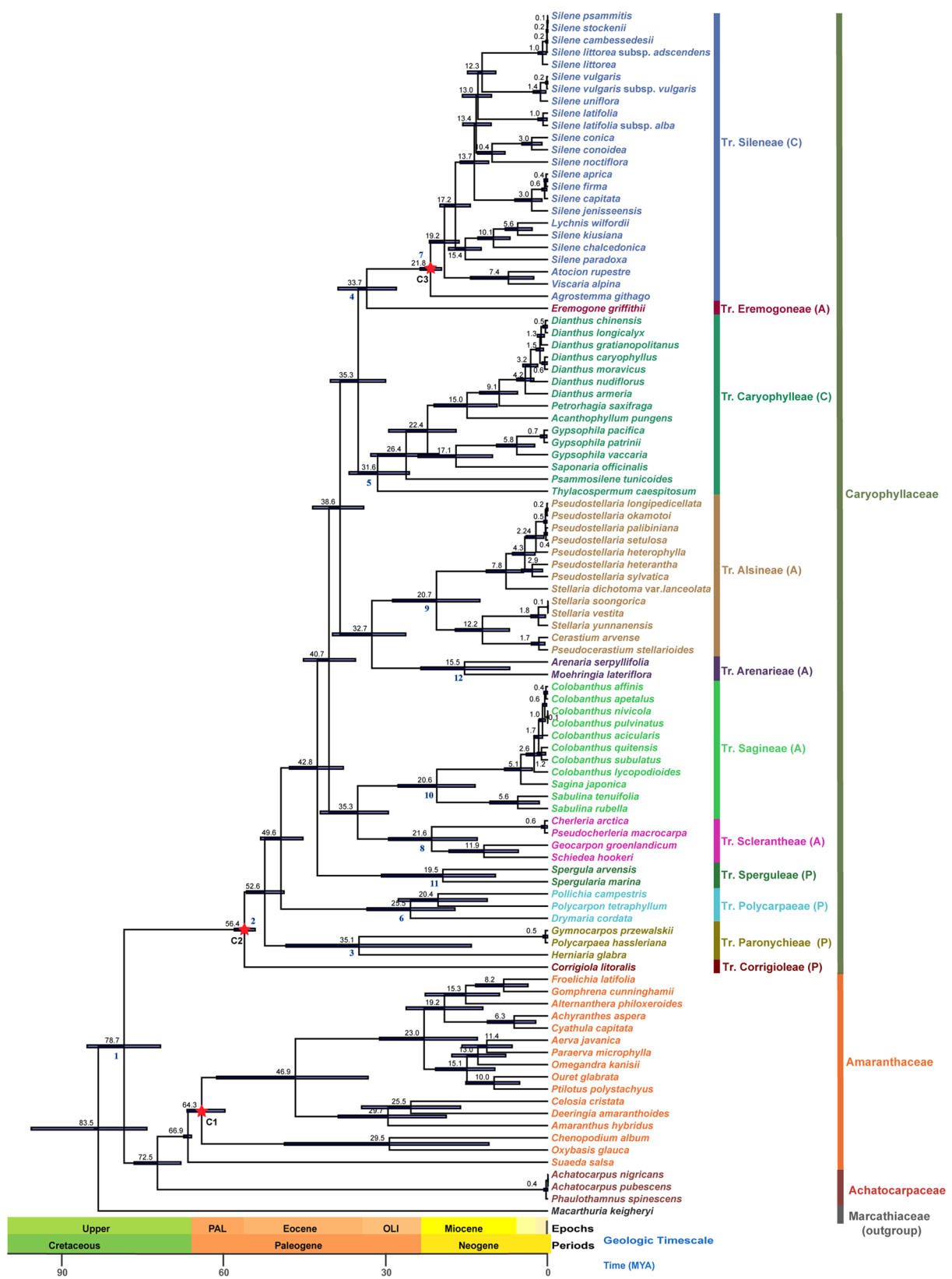


Fig. 4 Time-calibrated phylogeny of Caryophyllaceae based on the BEAST analysis of plastome-derived protein-coding (CDS) genes. Numbers in black circles indicate important divergence times within Caryophyllaceae. Red stars indicate calibration points (C1–C3); blue bars indicate 95% highest posterior density (HPD) intervals. Letters in parentheses show the assignment of each tribe to a traditionally recognized subfamily (A: Alsinoideae; C: Caryophylloideae; P: Paronychioideae). Time scale abbreviations: PAL, Paleocene; OLI, Oligocene

Table 5 Mean ages and 95% highest posterior density (HPD) intervals (in million years ago, Mya) of major nodes of the plastome-derived Caryophyllaceae phylogeny using BEAST analysis (see Fig. 4)

Node	Mean age, Mya	95% HPD, Mya
1	78.7	72.0–85.6
2	56.4	54.4–58.2
3	35.1	14.3–48.6
4	33.7	28.2–39.0
5	31.6	25.8–36.9
6	25.5	17.3–33.6
7	21.8	19.8–23.8
8	21.6	13.2–29.6
9	20.7	12.7–28.9
10	20.6	13.6–27.8
11	19.5	9.8–30.9
12	15.5	7.1–23.6

and dated its crown age to 24.55 Mya, hence close to our estimate (ca. 22.4 Mya). Using combined plastid markers (*rps16*, *matK*, *trnL-trnF*, *trnS-trnF*) data of nine tribes in Caryophyllaceae (excluding Corrigioleae and Paronychieae), Xu et al. [51] dated the crown age of Sileneae to 19 Mya, which again is consistent with our dating (ca. 21.8 Mya); however, their crown age for Caryophylleae (ca. 20 Mya) is later than our estimate (ca. 31.6 Mya). Compared with the above studies, Feng et al. [50] inference of the evolutionary history of Caryophyllaceae based on transcriptomics is more comprehensive. Their crown age for Caryophyllaceae (ca. 88.6 Mya) is earlier than our estimate (ca. 56.4 Mya); divergence times of each tribe are earlier than our estimates. Divergence time estimates may be sensitive to calibration choice and the sparse fossil record of Caryophyllaceae. The observed discrepancies may be attributed to the selection of fossil calibration points, differences in evolutionary rates between nuclear and plastid genomes, as well as potential deviations of plastid genome history from the species tree due to incomplete lineage sorting or hybridization. These factors together highlight the necessity for further investigation.

Based on our results, it is feasible that the initial divergence of Caryophyllaceae in the Late Paleocene (ca. 56.4 Mya) was consistent with prominent temperature increases during the Early Cenozoic (ca. 59–52 Mya; [57, 58]), including the Paleocene–Eocene Thermal Maximum (PETM; ca. 55.5 Mya; [59]). In fact, this climatic event has previously been implied in the diversification of many other groups of angiosperms, including Orchidaceae [60], Asteraceae [61], and Alzateaceae/Crypteroniaceae [62]. However, paleogeographic events may also have played a role in the initial but also subsequent divergence of Caryophyllaceae, including the emergence of both the North Atlantic (“Greenland”) and Bering

land bridges. These land bridges are well-known to have served as migration routes for temperate and (sub)tropical plants between North America and Eurasia, during most of the Cenozoic era (e.g. [63–67]). Notably, too, global temperatures dropped since the Early Eocene (ca. 52 Mya), and a large number of Caryophyllaceae tribes diverged during this period (e.g. Paronychieae, Polycarpaeae, and Sparguleae; stem ages: ca. 52.6–42.8 Mya; Fig. 4). We hypothesize that this may be possibly because these cooler temperatures increased the global heterogeneity of ecosystems and generated new habitats, promoting adaptive radiations, though the specific reasons require further investigation. (e.g. [68, 69].). Nevertheless, the majority of tribes (represented herein by multiple species) only started to diversify during the Late Eocene (i.e. Paronychieae, Caryophylleae; crown ages: ca. 35.1 and 31.6 Mya, respectively) and, in particular, during the Late Oligocene to Mid-Miocene (i.e. Polycarpaeae, Sparguleae, Scleranthaeae, Sagineae, Alsineae, Arenarieae, Sileneae; 25.5–15.5 Mya). Even though these tribal crown ages have broad HPD intervals (Fig. 4), indicating a degree of uncertainty, it can be speculated that the origin and diversification of many, if not the majority of Caryophyllaceae genera may be associated with global cooling and/or aridification events since the Mid-to-Late Miocene as well as the Pliocene–Pleistocene (e.g. *Silene*: [23]).

Character evolution of Caryophyllaceae

In this study, we subjected the plastome-derived MCC tree of Caryophyllaceae to BBM analysis to infer the evolution of five morphological traits distinguishing the three traditional subfamilies (Fig. 5). This analysis demonstrated that the MRCA of this family had stipules, free sepals, flowers apetalous with five stamens, and an utricle or achene fruit (as all retained in Corrigioleae, Paronychieae, and Polycarpaeae). Our results are consistent with those of Greenberg and Donoghue [4], who have also studied character evolution of Caryophyllaceae. Earlier Fior et al. [11] considered that dehiscent fruits (capsules) and stipules present were the ancestral characters of Caryophyllaceae, a view supported by our results. Stipules, which are lacking in Achatocarpaceae and Amaranthaceae [4], have been considered another ancestral characteristic of Caryophyllaceae [3]. Therefore, they might have originated independently in this family [4, 12]. However, stipules only exist in early-diverging groups, i.e. all species of Corrigioleae, Paronychieae, Polycarpaeae and most species of Sparguleae [3, 4, 12]. Notably, it has long been hypothesized that the ancestral flower condition of Caryophyllaceae was apetalous and petals were interpreted as centrifugally derived staminode appendages [4, 70–72]. However, combined with ontogenetic results, Wei and Ronse De Craene [73]

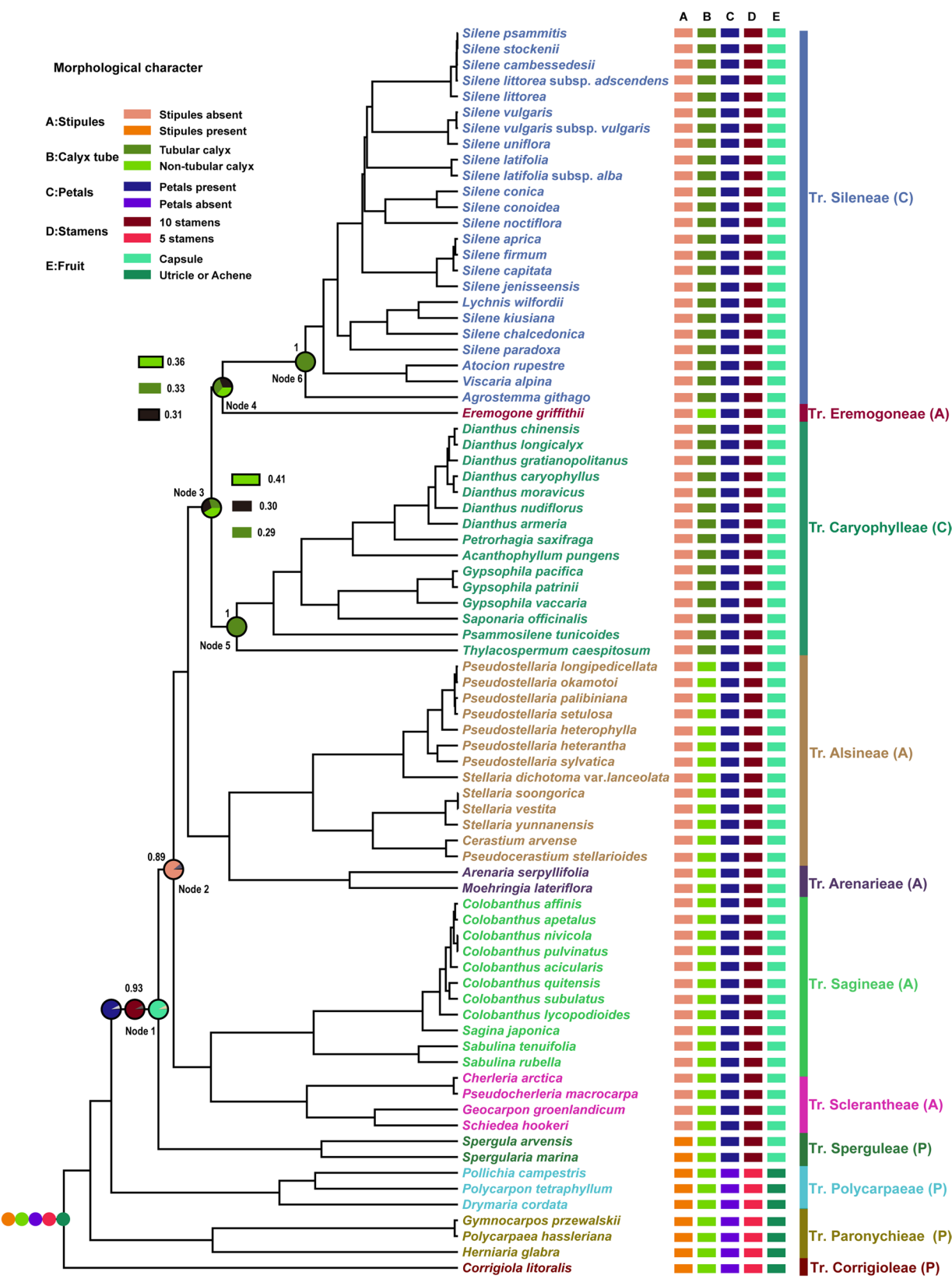


Fig. 5 Ancestral state reconstructions of five morphological characters along the BEAST-derived tree topology of Caryophyllaceae (see Fig. 4) using BBM (Bayesian Binary MCMC) analysis in RASP. A: stipules; B: calyx tube; C: petals; D: stamen number; E: fruit. Numbers at selected nodes (1–6) represent the relative probability (p) of occurrence of different ancestral states at these nodes. Letters in parentheses show the assignment of each tribe (Tr.) to a traditionally recognized subfamily (A: Alsinoideae; C: Caryophylloideae; P: Paronychioideae)

suggested that the ancestral condition of Caryophyllaceae was probably petaliferous, but petals disappeared in early diverging clades of Caryophyllaceae and reappeared in later derived clades. Our results also support this view. A tubular calyx has long been considered a synapomorphy of the traditional subfamily Caryophylloideae [3].

However, our BBM analysis suggests that a tubular calyx is either a symplesiomorphy that evolved in the MRCA of Caryophylleae–Eremogoneae–Sileneae and then was lost in Eremogoneae, or this trait represents a homoplasy that evolved twice in this group (in Caryophylleae and Sileneae, respectively). Considering the number of stamens, this is known to be highly variable in Caryophyllales and closely related to the number of petals [72]. According to our BBM reconstructions (Fig. 5), the MRCA of Caryophyllaceae had five stamens, as found in Amaranthaceae [74] and retained in Corrigioleae, Paronychieae, and Polycarpaeae. Interestingly, the number of stamens increased from five to 10 in the MRCA of the clade comprising Sperguleae and the remaining (seven) tribes of Caryophyllaceae, which is obviously related to the concomitant acquisition of petals in the ancestor of this clade [4]. It is important to acknowledge the inherent uncertainty in ancestral state reconstructions, especially for a family like Caryophyllaceae with a complex history potentially involving homoplasy and convergence [4, 8]. The reconstructions represent the most probable states under our model, but alternative scenarios exist, particularly at nodes with lower support. Thus, our conclusions should be viewed as a well-supported yet provisional hypothesis of trait evolution.

Finally, it is worth briefly reviewing additional findings of other studies on character evolution in Caryophyllaceae (i.e. for traits not examined herein). For example, Zhang et al. [17] analyzed the evolution of six traits in *Pseudostellaria* and related genera (i.e. habit; root tubers; cleistogamous flowers; petals, number of stamens and styles of chasmogamous flowers), concluding that the ancestral states of core *Pseudostellaria* include a tuberous root and cleistogamous flowers. Yao et al. [24] showed that the ancestral characteristics of Alsineae include three styles and six lobes at the apex of capsule. Finally, by studying character evolution in Caryophylleae, Madhani et al. [18] suggested that membranous commissures between the sepals may be an ancestral character in Caryophyllaceae. Their study further concluded that “most of the diagnostic traits used formerly in tribe Caryophylleae are homoplasious and not useful for defining the boundaries between the genera” [18].

Conclusions

This study is the first comprehensive phylogenomic analysis of Caryophyllaceae based on plastome-derived protein-coding (CDS) sequence data. Our strongly

supported ML, BI and time-calibrated (BEAST) phylogenies provide a solid foundation for further phylogenetic and biogeographic studies in this family. At this stage, our phylogenies strongly support the monophyly of Caryophyllaceae and date their crown age to the Late Paleocene (ca. 56.4 Mya). Although none of the three traditional subfamilies proved monophyletic, our plastome data fully resolved relationships among the 11 previously recognized tribes, with crown ages ranging from the Late Eocene to the Mid-Miocene (ca. 35.1–15.5 Mya). Ancestral trait reconstructions indicated that the ancestral states of Caryophyllaceae include: presence of stipules; free sepals; absence of petals; five stamens; and utricle or achene fruit. Notably, across the 79 Caryophyllaceae plastomes analysed, we also identified including 13 highly variable plastid regions, comprising one coding sequence (*infA*) and 12 intergenic spacer regions. These highly variable plastid regions will serve as powerful molecular markers for future phylogeographic, population genetic and/or conservation studies in Caryophyllaceae.

Acknowledgements

We sincerely thank the DNA bank of Missouri Botanical Garden for providing valuable research materials. We also gratefully acknowledge helpful comments made by two anonymous reviewers on an earlier draft of this manuscript.

Authors' contributions

Zhaoping Yang, Pan Li, and Haiwen Li designed this study; Pan Li contributed plant materials. Haiwen Li and Shiqiang Song performed the plastome assembly, annotation, and phylogenetic analyses. Pan Li identified habit characteristics. Haiwen Li performed the time calibration, diversification analysis, and ancestral state reconstruction. Haiwen Li, Zhaoping Yang and Hans Peter Comes wrote the manuscript. All authors have read and approved the final manuscript.

Funding

This research was supported by the National Natural Science Foundation of China (Grant Nos. 32400174, 32570239 and 32060053) and the Third Xinjiang Scientific Expedition Program (Grant No. 2022xjkk1505).

Data availability

All 33 plastome sequences have been uploaded to National Center for Biotechnology Information (NCBI).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 3 September 2025 / Accepted: 24 February 2026

Published online: 09 March 2026

References

1. Hernández-Ledesma P, Berendsohn WG, Borsch T, Von Mering S, Akhani H, Arias S, et al. A taxonomic backbone for the global synthesis of

- species diversity in the angiosperm order Caryophyllales. Willdenowia. 2015;45(3):281–383.
2. Yao G, Jin J-J, Li H-T, Yang J-B, Mandala VS, Croley M, et al. Plastid phylogenomic insights into the evolution of Caryophyllales. *Mol Phylogenet Evol*. 2019;134:74–86.
 3. Bittrich V. Caryophyllaceae. In: Kubitzki K, Rohrer J, Bittrich V, editors. The families and genera of vascular plant. Magnoliid, Hamamelid and Caryophyllid families. Volume 2. Berlin: Springer; 1993. pp. 206–36.
 4. Greenberg AK, Donoghue MJ. Molecular systematics and character evolution in Caryophyllaceae. *Taxon*. 2011;60(6):1637–52.
 5. Chandra S, Rawat D. Medicinal plants of the family Caryophyllaceae: a review of ethno-medicinal uses and pharmacological properties. *Integr Med Res*. 2015;4(3):123–31.
 6. Korkmaz M, Özçelik H. Economic importance of *Gypsophila* L., *Ankyropetalum fenzl* and *Saponaria* L. (Caryophyllaceae) taxa of Turkey. *Afr J Biotechnol*. 2011;10(47):9533–41.
 7. Kolodziej B, Okon S, Nucia A, Ociepa T, Luchowska K, Sugier D, et al. Morphological, chemical, and genetic diversity of *Gypsophila* L. (Caryophyllaceae) species and their potential use in the pharmaceutical industry. *Turkish J Bot*. 2018;42(3):257–70.
 8. Chrtěk J, Slavikova Z. Leitbündelanordnung in den Kronblättern von ausgewählten Arten der Familie Stellariaceae. *Preslia*. 1987;60:11–21.
 9. Fernandes A. Contribution à la connaissance cytotoxinomique des Spermatophyta du Portugal. II. Compositae. *Bol Soc Brot ser 2*. 1971;45:5–121.
 10. Smitsen RD, Clement JC, Garnock-Jones PJ, Chambers GK. Subfamilial relationships within Caryophyllaceae as inferred from 5' *ndhF* sequences. *Am J Bot*. 2002;89(8):1336–41.
 11. Fior S, Karis PO, Casazza G, Minuto L, Sala F. Molecular phylogeny of the Caryophyllaceae (Caryophyllales) inferred from chloroplast *matK* and nuclear rDNA ITS sequences. *Am J Bot*. 2006;93(3):399–411.
 12. Harbaugh DT, Nepokroeff M, Rabeler RK, McNeill J, Zimmer EA, Wagner WL. A new lineage-based tribal classification of the family Caryophyllaceae. *Int J Plant Sci*. 2010;171(2):185–98.
 13. Pusalkar PK. Redefining *Thylacospermum* and a New Tribe Thylacospermeae (Caryophyllaceae). *J Japanese Bot*. 2015;90(5):351–5.
 14. Dillenberger MS, Kaderit JW. Maximum polyphyly: Multiple origins and delimitation with plesiomorphic characters require a new circumscription of *Minuartia* (Caryophyllaceae). *Taxon*. 2014;63(1):64–88.
 15. Pirani A, Zarre S, Pfeil BE, Bertrand YJ, Assadi M, Oxelman B. Molecular phylogeny of *Acanthophyllum* (Caryophyllaceae: Caryophyllaceae), with emphasis on infrageneric classification. *Taxon*. 2014;63(3):592–607.
 16. Montesinos-Tubée DB, Borsch T. Molecular phylogenetics and morphology reveal the Plettkea lineage including several members of Arenaria and Pycnophyllopsis to be a clade of 21 South American species nested within Stellaria (Caryophyllaceae, Alsineae). *Willdenowia*. 2023;53(3):115–48.
 17. Zhang M-L, Zeng X-Q, Li C, Sanderson SC, Byalt VV, Lei Y. Molecular phylogenetic analysis and character evolution in *Pseudostellaria* (Caryophyllaceae) and description of a new genus, Hartmaniella, in North America. *Bot J Linn Soc*. 2017;184(4):444–56.
 18. Madhani H, Rabeler R, Pirani A, Oxelman B, Heubl G, Zarre S. Untangling phylogenetic patterns and taxonomic confusion in tribe Caryophyllaceae (Caryophyllaceae) with special focus on generic boundaries. *Taxon*. 2018;67(1):83–112.
 19. Fassou G, Korotkova N, Nersesyan A, Koch MA, Dimopoulos P, Borsch T. Taxonomy of *Dianthus* (Caryophyllaceae)—overall phylogenetic relationships and assessment of species diversity based on a first comprehensive checklist of the genus. *PhytoKeys*. 2022;196:91–214.
 20. Arabi Z, Ghahremaninejad F, Rabeler RK, Sokolova I, Weigend M, Zarre S. Inter-generic relationships within the tribe Alsineae (Caryophyllaceae) as inferred from nrDNA ITS and cpDNA *rps16* sequences: A step toward a phylogenetically based generic system. *Taxon*. 2022;71(3):608–29.
 21. Jafari F, Zarre S, Gholipour A, Eggens F, Rabeler RK, Oxelman B. A new taxonomic backbone for the infrageneric classification of the species-rich genus *Silene* (Caryophyllaceae). *Taxon*. 2020;69(2):337–68.
 22. Mahmoudi Shamsabad M, Moharrek F, Assadi M, Nieto Feliner G. Biogeographic history and diversification patterns in the Irano-Turanian genus *Acanthophyllum* s.l. (Caryophyllaceae). *Plant Biosystems-An Int J Dealing all Aspects Plant Biology*. 2021;155(3):425–35.
 23. Moilola NA, Mesbah M, Nylander S, Manning J, Forest F, de Boer HJ, et al. Biogeographic origins of southern African *Silene* (Caryophyllaceae). *Mol Phylogenet Evol*. 2021;162:107199.
 24. Yao G, Xue B, Liu K, Li Y, Huang J, Zhai J. Phylogenetic estimation and morphological evolution of Alsineae (Caryophyllaceae) shed new insight into the taxonomic status of the genus *Pseudocerasium*. *Plant Divers*. 2021;43(4):299–307.
 25. Zhou Y, Zheng Z. Genetic diversity and inter-relationship among *Stellaria* L. (Caryophyllaceae) species by ISSR markers. *Genetika*. 2022;54(1):119–30.
 26. Folk RA, Stubbs RL, Mort ME, Cellinese N, Allen JM, Soltis PS, et al. Rates of niche and phenotype evolution lag behind diversification in a temperate radiation. *Proc Natl Acad Sci*. 2019;116(22):10874–82.
 27. Soltis PS, Folk RA, Soltis DE. Darwin review: angiosperm phylogeny and evolutionary radiations. *Proc Royal Soc B*. 2019;286(1899):20190099.
 28. Dierckxens N, Mardulyn P, Smits G. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res*. 2017;45(4):e18.
 29. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28(12):1647–9.
 30. Huang DI, Cronk QC. Plann: a command-line application for annotating plastome sequences. *Appl plant Sci*. 2015;3(8):150026.
 31. Lohse M, Drechsel O, Kahlau S, Bock R. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res*. 2013;41(W1):W575–81.
 32. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol*. 2017;34(12):3299–302.
 33. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–80.
 34. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, et al. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol Ecol Resour*. 2020;20(1):348–55.
 35. Stamatakis A, RAxML-VI-HPC. Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006;22(21):2688–90.
 36. Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 2012;61(3):539–42.
 37. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: 2010 gateway computing environments workshop (GCE): 2010. IEEE; 1–8. <https://doi.org/10.1109/GCE.2010.5676129>
 38. Darriba D, Taboada G, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Meth*. 2012;9:772.
 39. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, et al. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comput Biol*. 2019;15(4):e1006650.
 40. Nowicke JW. Caryophyllales: Evolution and Systematics. In: Behnke HD, Mabry JT, editors. Pollen morphology and exine ultrastructure. Berlin: Springer; 1994. pp. 168–221.
 41. Punt W, Hoen PP. Caryophyllaceae. *Rev Palaeobot Palynol*. 1995;88:82–272.
 42. Jordan GJ, Macphail MK. A middle-late Eocene inflorescence of Caryophyllaceae from Tasmania, Australia. *Am J Bot*. 2003;90(5):761–8.
 43. Gizaw A, Brochmann C, Nemomissa S, Wondimu T, Masao CA, Tsuijime FM, et al. Colonization and diversification in the African 'sky islands': insights from fossil-calibrated molecular dating of *Lychnis* (Caryophyllaceae). *New Phytol*. 2016;211(2):719–34.
 44. Kaderit G, Borsch T, Weising K, Freitag H. Phylogeny of Amaranthaceae and Chenopodiaceae and the evolution of C₄ photosynthesis. *Int J Plant Sci*. 2003;164(6):959–86.
 45. Kaderit G, Gotzek D, Freitag H. Origin and age of Australian Chenopodiaceae. *Organisms Divers Evol*. 2005;5(1):59–80.
 46. Kaderit G, Ackerly D, Pirie MD. A broader model for C₄ photosynthesis evolution in plants inferred from the goosefoot family (Chenopodiaceae ss). *Proc Royal Soc B: Biol Sci*. 2012;279(1741):3304–11.
 47. Li H-T, Yi T-S, Gao L-M, Ma P-F, Zhang T, Yang J-B, et al. Origin of angiosperms and the puzzle of the Jurassic gap. *Nat plants*. 2019;5(5):461–70.
 48. Srivastava SK. Assorted angiosperm pollen from the Edmonton formation (Maestrichtian), Alberta, Canada. *Can J Bot*. 1969;47(6):975–89.
 49. Yu Y, Harris AJ, Blair C, He X. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Mol Phylogenet Evol*. 2015;87:46–9.

50. Feng K, Walker JF, Marx HE, Yang Y, Brockington SF, Moore MJ, et al. The link between ancient whole-genome duplications and cold adaptations in the Caryophyllaceae. *Am J Bot*. 2024;111(8):e16350.
51. Xu B, Luo D, Li ZM, Sun H. Evolutionary radiations of cushion plants on the Qinghai-Tibet Plateau: Insights from molecular phylogenetic analysis of two subgenera of *Arenaria* and *Thylacospermum* (Caryophyllaceae). *Taxon*. 2019;68(5):1003–20.
52. Stover L, Partridge A. Tertiary and Late Cretaceous spores and pollen from the Gippsland Basin, southeastern Australia. *Proc R Soc Victoria*. 1973;85(2):237–86.
53. Raine JI. Outline of a palynological zonation of Cretaceous to Paleogene terrestrial sediments in West Coast region, South Island, New Zealand. *New Z Geol Surv Rep*. 1984;109:1–82.
54. Truswell E, Sluiter I, Harris W. Palynology of the Oligocene-Miocene sequence in the Oakvale-1 corehole, western Murray Basin, South Australia. *BMR J Australian Geol Geophys*. 1985;9:267–95.
55. Chandler MEJ. Flora of the Lower Headon beds of Hampshire and the Isle of Wight. *Bull Br Museum Nat History: Geol*. 1961;5:91–158.
56. Frajman B, Eggens F, Oxelman B. Hybrid origins and homoploid reticulate evolution within *Heliosperma* (Sileneae, Caryophyllaceae)—a multigene phylogenetic approach with relative dating. *Syst Biol*. 2009;58(3):328–45.
57. Zachos J, Pagani M, Sloan L, Thomas E, Billups K. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science*. 2001;292(5517):686–93.
58. Zachos JC, Dickens GR, Zeebe RE. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature*. 2008;451(7176):279–83.
59. Bowen GJ, Maibauer BJ, Kraus MJ, Röhl U, Westerhold T, Steimke A, et al. Two massive, rapid releases of carbon during the onset of the Palaeocene–Eocene thermal maximum. *Nat Geosci*. 2015;8(1):44–7.
60. Ramírez SR, Gravendeel B, Singer RB, Marshall CR, Pierce NE. Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. *Nature*. 2007;448(7157):1042–5.
61. Barreda VD, Palazzesi L, Tellería MC, Olivero EB, Raine JI, Forest F. Early evolution of the angiosperm clade Asteraceae in the Cretaceous of Antarctica. *Proceedings of the National Academy of Sciences*. 2015;112(35):10989–10994.
62. Morley RJ. Interplate dispersal paths for megathermal angiosperms. *Perspect Plant Ecol Evol Syst*. 2003;6(1–2):5–20.
63. Wen J. Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. *Annu Rev Ecol Syst*. 1999;30(1):421–55.
64. Graham A. The role of land bridges, ancient environments, and migrations in the assembly of the North American flora. *J Syst Evol*. 2018;56(5):405–29.
65. Wen J, Nie ZL, Ickert-Bond SM. Intercontinental disjunctions between eastern Asia and western North America in vascular plants highlight the biogeographic importance of the Bering land bridge from late Cretaceous to Neogene. *J Syst Evol*. 2016;54(5):469–90.
66. Grimsen F, Denk T, Zetter R. Cenozoic Vegetation and Phytogeography of the Sub-arctic North Atlantic. In: *Biogeography in the Sub-Arctic*. 2021; 29–49. <https://doi.org/10.1002/9781118561461.ch2>
67. Jiang D, Klaus S, Zhang Y-P, Hillis DM, Li J-T. Asymmetric biotic interchange across the Bering land bridge between Eurasia and North America. *Natl Sci Rev*. 2019;6(4):739–45.
68. Haffer Jr. Alternative models of vertebrate speciation in Amazonia: an overview. *Biodivers Conserv*. 1997;6:451–76.
69. Gustafsson ALS, Verola CF, Antonelli A. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus *Hoffmannseggella* (Orchidaceae: Epidendroideae). *BMC Evol Biol*. 2010;10:1–13.
70. De Ronse LP. Homology and evolution of petals in the core eudicots. *Syst Bot*. 2008;33(2):301–25.
71. Brockington SF, Alexandre R, Ramdial J, Moore MJ, Crawley S, Dhirga A, et al. Phylogeny of the Caryophyllales *sensu lato*: revisiting hypotheses on pollination biology and perianth differentiation in the core Caryophyllales. *Int J Plant Sci*. 2009;170(5):627–43.
72. De Ronse LP. Reevaluation of the perianth and androecium in Caryophyllales: implications for flower evolution. *Plant Syst Evol*. 2013;299:1599–636.
73. Wei L, Craene LRD. What is the nature of petals in Caryophyllaceae? Developmental evidence clarifies their evolutionary origin. *Ann Botany*. 2019;124(2):281–95.
74. Townsend CC. *Amaranthaceae*. In: *The families and genera of vascular plants*. Edited by Kubitzki J. Springer-Verlag: Berlin. 1993;2:70–91.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.