

Comparative plastome genomics of *Lobelia* in the Mexican Transition Zone reveals conserved genome structure (Campanulaceae: Lobelioideae)

Miguel A. Pérez-Pérez

map758@nau.edu

Northern Arizona University

Kyle G. Keepers

University of Colorado

Nolan C. Kane

University of Colorado

Arturo Castro-Castro

Jardín Etnobiológico Estatal de Durango, Unidad – Consejo Nacional de Ciencia y Tecnología

Richard G. J. Hodel

Northern Arizona University

Laura P. Lagomarsino

Louisiana State University

Tina J Ayers

Northern Arizona University

Research Article

Keywords: Chloroplast genome, infA gene, Inversions, Phylogenomics, Structural rearrangements

Posted Date: October 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11089280/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Plastid genomes are generally conserved in angiosperms; however, structural modifications including inversions, gene losses, and shifts in inverted repeat (IR) boundaries are widespread across some lineages. Campanulaceae is particularly notable for extensive plastome rearrangements, including inversions, gene loss and pseudogenization, repeat-associated structural changes, and shifts in IR boundaries. Comparative plastome studies of the species-rich genus *Lobelia* remain limited, and the five infrageneric sections represented in the Mexican Transition Zone (MTZ) have not previously been investigated within a comparative plastome framework.

Results

We assembled and annotated 41 plastomes, comprising 38 *Lobelia* representing the five sections occurring in the MTZ and three *Diastatea*. Overall plastome organization was highly conserved, with limited structural variation. *Lobelia gypsophila* exhibited a large inversion in the large single-copy (LSC) region, whereas *L. yucatana* showed a reduction in IR length associated with a distinct IR/SSC boundary configuration. Among the 41 newly assembled plastomes, *infA* was annotated as a CDS in 30 taxa, occurred as a pseudogenized copy in five, and was absent in six. Phylogenetic analyses recovered sects. *Cryptostemon*, *Homochilus*, *Lobelia*, and *Tylomium* as monophyletic, whereas sect. *Stenotium* was polyphyletic and distributed across four lineages. The three sampled species of *Diastatea* were not recovered as monophyletic and occurred within a clade that also included five species assigned to *Lobelia* sect. *Stenotium*.

Conclusions

Plastome evolution among *Lobelia* and *Diastatea* of the MTZ was comparatively conservative despite the extensive structural rearrangements documented more broadly across Campanulaceae. The principal structural differences were restricted to individual species, with an LSC inversion in *L. gypsophila* and an IR reduction in *L. yucatana*, while the distribution of *infA* showed additional variation among lineages. Plastome phylogenetic relationships were largely congruent with recent nuclear phylogenomic results and further support the non-monophyly of sect. *Stenotium* under its current circumscription.

Introduction

Chloroplast genomes (plastomes) of land plants are generally characterized by conserved gene content, gene order and genome structure over hundreds of millions of years of evolution [1]. The conserved structure and content are likely due to plastomes' fundamental role in photosynthesis and other essential metabolic processes essential for plant growth and stress response [2]. Plastid genomes are typically inherited uniparentally, most often through the maternal lineage in angiosperms [3], although paternal inheritance has been documented in some gymnosperms (e.g., genus *Sequoia*) [4]. Structurally, plastomes exhibit a quadripartite organization consisting of two identical copies of the inverted repeat (IR) regions that separate the large single-copy (LSC) and the small single-copy (SSC) regions. Plastomes typically range in size from 120 to 170 kb in length and encode approximately 110–130 genes (~ 80 proteins, ~ 30 tRNAs, and four rRNAs) [5]. Despite their overall conservation, plastomes harbor phylogenetically informative variation that has been widely used to resolve evolutionary relationships across a broad range of taxonomic levels [6, 7].

Plastomes are generally conserved, although numerous angiosperm lineages have experienced substantial structural modifications. Gene losses, inversions, deletions, intron losses, expansion or contraction of inverted repeats, and, in rare cases, complete loss of an IR region have all contributed to plastome diversification [8–13]. Large inversions are comparatively uncommon but have been documented in several plant groups, including Poaceae [14] and Asteraceae [15]. Likewise, variation in IR boundaries represents an important mechanism underlying plastome size evolution. For example, extensive IR expansion and contraction have been reported in more than 30 species of *Angelica* (Apiaceae) [16]. Collectively, these structural modifications can alter plastome synteny (gene order) and provide valuable insights into the rates of genome evolution and the regulation of gene expression.

Among angiosperms, Campanulaceae is notable for exhibiting substantial plastome structural variation, including inversions, gene losses (e.g. plastid-encoded *accD* gene, *clpP*, *infA*, and *rpL23*), shifts in inverted repeat boundaries, and variation in genome size [17–25]. Additional modifications reported across the family include the loss of the *rps16* intron, pseudogenization of *infA* and *rpL23*, and reduction of the SSC region associated with IR expansion [20]. An approximately 11-kb expansion of the IR into the SSC has been proposed as an early event during plastome evolution within Campanulaceae, representing a major structural change in plastome organization [23]. Structural rearrangements, gene duplications, and repetitive DNA elements are also thought to contribute to plastome size variation, as demonstrated

in *Campanula americana* [24]. Because plastomes combine highly conserved gene content with lineage-specific structural rearrangements, they provide an excellent system for investigating genome evolution while simultaneously reconstructing phylogenetic relationships.

Within Campanulaceae, *Lobelia* is the largest and most diverse genus of the subfamily Lobelioideae [26]. However, molecular phylogenetic studies have demonstrated that *Lobelia*, as traditionally circumscribed, is polyphyletic [26–29]. Nevertheless, the infrageneric classification proposed by Lammers [30], which recognizes 18 sections of *Lobelia* based primarily on morphological characters, remains a useful framework for discussing diversity and evaluating sectional relationships.

Lobelia is cosmopolitan but is particularly species-rich in tropical Africa, southern Africa, and Mexico, which represent major centers of diversity [31]. Mexico harbors approximately 85 recognized species, representing 19% of the approximately 443 species currently recognized worldwide in the World Flora Online Plant List (<https://wfoplantlist.org>). Following the infrageneric classification of Lammers [30], these species belong to five of the eighteen recognized sections: sect. *Cryptostemon*, sect. *Homochilus*, sect. *Lobelia*, sect. *Tylomium* (all restricted to the Americas), and sect. *Stenotium* (cosmopolitan). Together, these sections represent one of the largest and most diverse assemblages traditionally assigned to *Lobelia*. Most species belonging to these five sections occur within the Mexican Transition Zone (MTZ), a biogeographically complex region where the Nearctic and Neotropical floras converge and one of the principal centers of plant diversification in North America [32]. Although the MTZ has received considerable attention from floristic and evolutionary perspectives [33, 34], comparatively little attention has been devoted to genomic studies of its flora. Recent phylogenomic analyses have substantially improved our understanding of evolutionary relationships among *Lobelia* lineages occurring in the MTZ [35]. Nevertheless, plastome evolution and structural variation across representatives of the five sections have not previously been investigated within a comparative plastome framework. Consequently, this assemblage provides an excellent opportunity to investigate plastome structural variation, genome evolution, and plastid phylogenetic relationships within one of the most diverse evolutionary assemblages traditionally assigned to *Lobelia*.

Comparative plastome studies in *Lobelia* remain scarce. Early investigations by Knox [36] characterized major plastome rearrangements in Lobeliaceae using restriction mapping, documenting the extensive IR expansion and multiple inversions that distinguish the family from the ancestral angiosperm plastome organization [36]. More recently, complete plastome comparisons have been conducted for species representing five of the eighteen sections recognized by Lammers [30] from East Asia [25]. That study identified structural rearrangements in sect. *Delostemon* and sect. *Stenotium* and documented substantial variation in IR length, ranging from 36,955 to 42,089 bp among sampled taxa [25]. Collectively, these studies indicate that plastome structural diversity extends across multiple *Lobelia* lineages while also highlighting the limited taxonomic representation available for comparative analyses.

Previous phylogenetic studies of Lobelioideae have relied primarily on single- or multi-locus DNA sequence datasets [27, 28, 37–41]. Advances in high-throughput sequencing have enabled the routine assembly of complete plastid genomes, providing substantially greater genomic information for comparative evolutionary and phylogenetic analyses. Consequently, complete plastomes have increasingly been used to investigate evolutionary relationships and biogeographic history in specific *Lobelia* lineages, such as the giant lobelias (sect. *Rhynchopetalum*) [42, 43]. However, no comparative plastome study has included numerous representatives of all five *Lobelia* sections occurring in the Mexican Transition Zone, leaving plastome diversity across one of the largest assemblages traditionally assigned to *Lobelia* largely unexplored.

Here, we use extensive sampling of *Lobelia* from the Mexican Transition Zone (MTZ), representing the five sections occurring in the region, together with the closely related genus *Diastatea*, to: (1) assemble and annotate new plastomes; (2) compare and characterize plastome organization and structural variation across species; and (3) reconstruct plastome phylogenetic relationships. By integrating comparative plastome genomics and phylogenetic analyses, this study provides new insights into chloroplast genome evolution in *Lobelia* and expands genomic resources for one of the most diverse lineages of Campanulaceae.

Methods

Taxon sampling

Plant material was obtained from herbarium sheets at the following institutions: Northern Arizona University (ASC), Universidad Nacional Autónoma de México (MEXU), Instituto de Ecología, A. C., Pátzcuaro (IEB), Instituto de Ecología, A. C., Xalapa (XAL), Instituto Politécnico Nacional, CIIDIR Unidad Durango (CIIDIR), Instituto Politécnico Nacional, (CIIDIR-Oax., I.P.N.) (OAX), Missouri Botanical Garden (MO) and Billie L. Turner Plant Resources Center (TEX-LL). The taxonomic determinations of all plant material used for DNA in this study were reviewed and confirmed by Tina J. Ayers, an expert in Lobelioideae. Full collector names and voucher information, including collector numbers and the institutional acronym of the herbarium where each specimen is housed, are provided in Table S1. Herbarium abbreviations follow the index Herbariorum (<https://sweetgum.nybg.org/science/ih/>).

Ingroup sampling, informed by a recent phylogenomic analysis [35], included three species of *Lobelia* sect. *Cryptostemon*, four species of *L.* sect. *Homochilus*, two species of *L.* sect. *Lobelia*, plus a downloaded plastome of *L. inflata* (NC_033368), 27 species of *L.* sect. *Stenotium*, and two species of *L.* sect. *Tylomium*, as well as three species of *Diastatea* subg. *Diastatea* [40] (Table S1), which is phylogenetically nested with sect. *Stenotium* within the MTZ clade [35]. Outgroups included two species of *Lobelia* sect. *Tupa* from our sampling, together with five additional Campanulaceae plastomes obtained from NCBI: *Cyphia angustiloba* (NC_036086), *C. tortilis* (NC_036094), *Campanula punctata* (NC_033337), *C. takesimana* (MW_013763), and *Adenophora erecta* (NC_036222).

DNA extraction and sequencing

For each sample, 20 mg of leaf tissue was ground using a Geno/Grinder Homogenizer. DNA extractions were performed using a single sucrose prewash plus-CTAB protocol [44]. We assessed the quality of the DNA with 1% agarose gels and quantified it with a Qubit™. Samples were submitted to RAPID Genomics (Gainesville, FL, USA) for assessment, quantification, library preparation, and sequencing. Whole genome sequencing (genome skimming) was performed on the Illumina NovaSeq 6000 platform (Illumina, San Diego, California, USA) using the S4 flow cell and with paired-end 250-bp read output.

Plastome assembly and annotation

Genomic read adapters were trimmed using fastp [45], and read quality was assessed with FastQC v 0.11.9 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) prior to de novo assembly. Trimmed reads from each sample were assembled de novo using GetOrganelle v1.7.6 [46] with the following parameters: *get_organelle_from_reads.py*, -1, r1_path, -2, r2_path, -t, 8, -o, output, -F, embplant_pt, -R, 15, -k, 21,45,65,85,95,105. The circular contig corresponding to pattern 1 was selected when recovered. Moreover, assembly graphs (GFA files), were visualized and edited in Bandage v0.8.1 [47] to confirm complete plastome paths and, when necessary, to manually resolve the assembly graph into a single complete contig.

After obtaining complete plastome sequences in FASTA format, each plastome was aligned against the previously published plastome of *Lobelia laxiflora* (NC_033370) using BLAST at NCBI. *Lobelia laxiflora*, a member of the study group, was used solely as a reference to standardize plastome orientation and coordinate position, rather than to infer the direction of structural evolution. Plastomes were oriented to begin at the start of the large single-copy (LSC) region, ensuring consistent representation of the LSC, small single-copy (SSC), and inverted repeat (IR) regions across taxa.

Genome annotation was performed using GeSeq [48] in CHLOROBX (<https://chlorobox.mpimp-golm.mpg.de/index.html>). We used the following plastomes as references: *Lobelia laxiflora* (NC_033370.1); *L. erinus* (NC_036098); *L. polyphylla* (NC_033373); *L. spicata* (NC_036080); *L. cardinalis* (NC_036097); *L. puberula* (NC_036078); *L. inflata* (NC_033368), and *Diastatea micrantha* (NC_035358), specified as “3rd Party References” in CHLOROBX. Circular plastome maps were generated using OGDRAW [49]. Annotations were summarized using a customized script to get the number and length of each coding gene. Moreover, annotations were manually inspected and adjusted in Geneious Prime→ 2025.0.3 [50], when necessary to correct gene boundaries and add genes that were missed by the first-pass annotation provided by GeSeq [48], as indicated by the gene inventory (Table S2). A final scan for open reading frames (ORFs) was performed using the ORF scanner function in Geneious Prime→ 2025.0.3 [50]. Newly annotated plastomes were submitted to NCBI GenBank under accession numbers PZ862060-PZ862100 (Table 1).

Table 1
General plastome information.

Plastome	Length_bp	GC_percent	LSC_bp	SSC_bp	IR_bp	CDS	tRNA	GenBank accession
<i>Diastatea_maximiliana</i> _TJA1774	165152	39.39	83741	7817	36797	90	37	PZ862060
<i>Diastatea_micrantha</i> _TJA1761	165355	39.53	82364	8035	37478	91	37	PZ862061
<i>Diastatea_virgata</i> _TJA1772	165827	39.64	83271	7986	37285	90	37	PZ862062
<i>Lobelia_alanae</i> _MJC31421	165988	39.17	86948	8026	35507	90	39	PZ862063
<i>Lobelia_anatina</i> _TJA1622	165751	39.19	83095	7824	37416	89	40	PZ862064
<i>Lobelia_berlandieri</i> _YRA843	166245	39.11	83261	7582	37701	89	37	PZ862065
<i>Lobelia_bryophila</i> _var_ <i>fimbriata</i> _VWS5660	166141	39.19	83481	8060	37300	89	38	PZ862066
<i>Lobelia_cardinalis</i> _KMH781	163654	39.93	84017	7957	35840	88	38	PZ862067
<i>Lobelia_cordifolia</i> _TRVD2007-1207	166183	39.25	83075	8098	37505	86	34	PZ862068
<i>Lobelia_dielsiana</i> _AVG3902	165911	39.2	83193	7978	37370	88	38	PZ862069
<i>Lobelia_divaricata</i> _DLP1579	165810	39.17	82504	7714	37796	90	36	PZ862070
<i>Lobelia_ehrenbergii</i> _var_ <i>ehrenbergii</i> _TJA1753	165760	39.12	82880	7858	37511	88	41	PZ862071
<i>Lobelia_excelsa</i> _TGL7604	166240	39.08	83648	8080	37256	90	38	PZ862072
<i>Lobelia_flexuosa</i> _var_ <i>flexuosa</i> _SG2341	166903	39.17	82692	8057	38077	88	38	PZ862073
<i>Lobelia_ghiesbreghtii</i> _RG3015	165533	39.14	83015	8008	37255	90	38	PZ862074
<i>Lobelia_goldmannii</i> _SAS235	167611	39.23	83432	8059	38060	89	38	PZ862075
<i>Lobelia_gruina</i> _MJC30278	165340	39.12	82989	7849	37251	89	41	PZ862076
<i>Lobelia_guerrerensis</i> _VWS5316	166562	39.22	83742	8034	37393	89	38	PZ862077
<i>Lobelia_gypsophilla</i> _TJA1752	168277	38.87	85156	7715	37703	90	37	PZ862078
<i>Lobelia_hartwegii</i> _VWS5918	167599	39.16	83841	8000	37879	90	39	PZ862079
<i>Lobelia_irasuensis</i> _var_ <i>fucata</i> _LLP350	166669	39.18	83186	7915	37784	91	38	PZ862080
<i>Lobelia_jaliscensis</i> _TJA1776	166589	39.25	83329	8012	37624	88	39	PZ862081
<i>Lobelia_knoblochii</i> _GR5574	167436	39.24	83872	8056	37754	90	37	PZ862082
<i>Lobelia_laxiflora</i> _var_ <i>angustifolia</i> _TRVD658	167355	39.21	84015	8042	37649	88	38	PZ862083
<i>Lobelia_laxiflora</i> _RGsn	166398	39.22	83566	8042	37395	89	38	PZ862084
<i>Lobelia_lobelioides</i> _RCD2061	166855	39.16	83816	7947	37546	89	37	PZ862085
<i>Lobelia_longicaulis</i> _JIC21345	165576	39.25	83135	8103	37169	87	38	PZ862086
<i>Lobelia_macrocentron</i> _HAG324	166662	39.2	83676	8066	37460	90	37	PZ862087
<i>Lobelia_margarita</i> _TJA1759	166226	39.07	83354	7684	37594	90	37	PZ862088
<i>Lobelia_mcvaughii</i> _TJA348	167721	39.25	83804	8063	37927	89	37	PZ862089
<i>Lobelia_modesta</i> _TJA1350	164484	39.21	82820	7820	36922	88	39	PZ862090
<i>Lobelia_nana</i> _CC72	164684	39.09	82869	7807	37004	89	38	PZ862091
<i>Lobelia_plebeia</i> _EM19646	165743	39.22	83634	8059	37025	87	38	PZ862092
<i>Lobelia_polyphylla</i> _TGL7723	166845	39.08	83835	8056	37477	90	38	PZ862093
<i>Lobelia_pulchella</i> _RGsn	165705	39.21	83059	8060	37293	87	38	PZ862094
<i>Lobelia_saliciana</i> _CLJ164	165092	38.93	82712	7864	37258	90	37	PZ862095

Plastome	Length_bp	GC_percent	LSC_bp	SSC_bp	IR_bp	CDS	tRNA	GenBank accession
<i>Lobelia_sartorii</i> _RG3006	164892	39.16	82751	7807	37167	90	38	PZ862096
<i>Lobelia_tatea</i> _RG3003	165814	38.93	82826	7864	37562	88	38	PZ862097
<i>Lobelia_volcanica</i> _VWS4737	166876	39.21	82909	8033	37967	89	37	PZ862098
<i>Lobelia_xalapensis</i> _TJA1790	164890	39.21	82954	7390	37273	89	37	PZ862099
<i>Lobelia_yucatanica</i> _GAM4737	160815	39.58	83021	12230	32782	85	39	PZ862100

Comparative genomics of whole plastomes

We removed one copy of the inverted repeat (IR) to avoid redundancy and aligned the 41 newly assembled plastomes using the progressive Mauve algorithm [51]. We performed a multiple sequence alignment to detect possible plastome rearrangements such as gene loss, reduction of IR, and inversions. To verify the inversion of *L. gypsophila*, we calculated the read-alignment depth across the whole plastome using BWA [52] for read alignment and SAMtools [53] to generate a sorted BAM file, followed by *samtools depth* to calculate read depth.

Locally collinear blocks (LCBs) identified by Mauve [51] were used to assess synteny and to detect rearranged regions across plastomes. We visualized the IR regions and their bordering regions using CPJSDraw software [54]. The IR/LSC and IR/SSC junctions were examined in the output visualizations of CPJSDraw to identify expansions and contractions of the IR regions among taxa. Genome size, gene content, and gene order were compared across plastomes to characterize variation in plastome organization. The package CPStools [55] was used to obtain the simple sequence repeats (SSR).

Phylogenetic analysis

We used two complementary datasets, whole-plastome sequences and shared protein-coding genes (CDS), to infer phylogenetic relationships among *Lobelia* and allied taxa. The analyses included the 41 newly assembled plastomes and six plastomes obtained from NCBI, resulting in a total dataset of 47 plastomes.

For the whole-plastome analysis, sequences were standardized to retain a single copy of the inverted repeat (IR), thereby avoiding duplication of the phylogenetic signal from the two IR copies, and were represented in LSC–IR–SSC configuration. Because preliminary conventional whole-plastome alignments contained extensive gap-rich regions associated with differences in genome organization among taxa, homologous regions were identified using progressive Mauve v2.4.0, snapshot build 13 February 2015 [51]. This approach allowed locally collinear blocks (LCBs) to be identified despite differences in their relative position or orientation resulting from structural rearrangements. The progressive Mauve XMFA output was parsed to identify core LCBs containing aligned nucleotide sequence from all 47 taxa; taxa without sequence represented in a given LCB were considered absent from that block. Thirty-three core LCBs meeting this criterion were retained for subsequent analyses, representing approximately 128,351 bp of homologous plastome sequence per taxon on average.

For each core LCB, Mauve alignment gaps were removed and nucleotide sequences were extracted for all taxa in the orientation represented in the XMFA. The 33 LCBs were then independently realigned using MAFFT v7.745 [56] with the `--auto` option and eight threads. The resulting alignments were concatenated using AMAS v1.0 [57], producing a 184,369-bp matrix, with LCB boundaries retained as partitions. Poorly aligned positions were removed using trimAl [58] with the `-automated1` heuristic, resulting in a final alignment of 120,798 nucleotide positions.

Nucleotide substitution-model selection for the whole-plastome matrix was performed with ModelTest-NG [58], and GTR + I+G4, selected according to AIC and AICc, was used for maximum-likelihood inference. The phylogeny was reconstructed in IQ-TREE v2.4.0 [59] using eight threads, with branch support assessed using 1,000 ultrafast bootstrap replicates and 1,000 SH-aLRT replicates.

For the CDS analysis, the DNA sequence of all annotated protein-coding genes were extracted from the same 47 plastomes using a custom Python script. Genes were compared across taxa to identify a common set of protein-coding loci, resulting in 67 CDS shared by all 47 plastomes. Each CDS was aligned independently using MAFFT v7.745 [59] with the `--auto` option, and the resulting alignments were manually inspected in Geneious Prime→ 2025.0.3 [50]. The 67 individual gene alignments were concatenated into a single supermatrix using AMAS v1.0 [57], while retaining gene boundaries as initial partitions.

Maximum-likelihood analysis of the concatenated CDS matrix was conducted in IQ-TREE v2.4.0 [59]. The MFP+MERGE procedure was used to select the best-fitting nucleotide substitution models and identify an optimized partitioning scheme by merging compatible gene partitions. Maximum-likelihood phylogenetic inference was then performed under the selected partitioning scheme, with branch support evaluated using 1,000 ultrafast bootstrap replicates.

Results

General results

Phylogenetic analyses resolved the sampled *Lobelia* and allied taxa into 11 well-supported major clades, providing a framework for evaluating patterns of plastome variation across the dataset. These clades are used throughout the following sections to describe and compare plastome characteristics among lineages. Overall, plastome structure and gene content were broadly conserved, with the main patterns of genomic and phylogenetic variation presented below.

Plastome general characteristics

All newly assembled plastomes exhibited the typical quadripartite structure consisting of a large single-copy (LSC) region, a small single-copy (SSC) region, and two inverted repeat (IR) regions. Plastome length ranged from 160 to 168 kb, with an average size of approximately 165 kb and a mean GC content of 39.3% (Table 1). Genome size varied moderately among sections – sect. *Cryptostemon* ranged from 166–167 kb, sect. *Homochilus* from 165–167 kb, sect. *Lobelia* from 163–166 kb, sect. *Tylomium* from 165–166 kb, and sect. *Stenotium* (identified to be polyphyletic, forming four separate subclades) from 160–168 kb (Table 1). Although sect. *Stenotium* was recovered as polyphyletic in the phylogenetic analyses, its constituent lineages showed broadly similar plastome sizes. A representative plastome map for *Lobelia gypsophila* in sect. *Stenotium* is shown in Fig. 1.

Comparative genomics

Whole-plastome comparisons revealed a high degree of collinearity across the sampled taxa, indicating that overall plastome organization was highly conserved (Fig. 2). The principal exception was *L. gypsophila* (sect. *Stenotium*), which contained a large inversion within the LSC region. The inverted region encompassed a contiguous genomic segment containing genes including *petA*, *psbC*, *psaA*, *rps4*, *ndhK*, and *rpoB* (Fig. 2). No additional large-scale rearrangements were detected among the remaining plastomes. The read alignment depth map for the *L. gypsophila* plastome is visualized in Figure S1, with arrows indicating the boundaries of the 24,412 bp inversion. Reads align cleanly across the inversion boundaries, with no anomalous spikes or dips in coverage.

Inverted repeat boundaries

The positions of the LSC/IR and SSC/IR junctions were highly conserved across most sampled plastomes (Fig. 3). The LSC/IRb junction was consistently located adjacent to *trnH* and *rps19*. The IRb/SSC and SSC/IRa junctions were associated with *ndhE*. The IRa/LSC junction was consistently associated with *trnH*-GUG, indicating limited variation in IR boundary positions across the dataset.

Lobelia yucatana (sect. *Stenotium*) represented the principal exception to this conserved pattern. Its IR region was reduced to 32,782 bp, the shortest among all sampled plastomes, whereas its SSC region reached 12,230 bp, the longest observed in the dataset (Table 1). *Lobelia yucatana* also possessed the smallest plastome analyzed, with a total length of 160,815 bp. In contrast, IR and SSC lengths varied only moderately among the remaining taxa, with no comparable shifts in IR boundary configuration.

Variation in *ycf1* annotation was detected among the plastomes examined. Most taxa contained two annotated copies of *ycf1*. In *Lobelia flexuosa* subsp. *flexuosa*, both copies were annotated as pseudogenes due to internal stop codons disrupting the open reading frame, whereas *L. cordifolia* differed from the predominant pattern by containing a single intact copy of *ycf1* located within IRa (Table 2).

Table 2
Distribution of *ycf1* within the inverted repeats of *Lobelia* and *Diastatea* plastomes from the Mexican Transition Zone.

Taxon	Voucher	<i>ycf1</i> in IRB (bp)	<i>ycf1</i> in IRA (bp)
<i>Diastatea maximiliana</i>	TJA1774	107,729–114,553	134,341–141,165
<i>Diastatea micrantha</i>	TJA1761	106,696–113,853	133,867–141,024
<i>Diastatea virgata</i>	TJA1772	107,522–114,571	134,528–141,577
<i>Lobelia alanae</i>	MJC31421	109,368–116,471	136,466–143,563
<i>Lobelia anatina</i>	TJA1622	107,517–114,545	134,302–141,330
<i>Lobelia berlandieri</i>	YRA843	107,505–114,980	134,527–142,002
<i>Lobelia bryophila</i> var. <i>fimbriata</i>	VWS5660	107,719–114,786	134,837–141,904
<i>Lobelia cardinalis</i>	KMH781	106,928–113,887	133,785–140,744
<i>Lobelia cordifolia</i>	TRVD2007-1207	—	134,659–141,738
<i>Lobelia dielsiana</i>	AVG3902	107,479–114,576	134,529–141,626
<i>Lobelia divaricata</i>	DLP1579	107,061–114,308	134,007–141,254
<i>Lobelia ehrenbergii</i> subsp. <i>ehrenbergii</i>	TJA1753	106,785–113,819	133,612–140,646
<i>Lobelia flexuosa</i> subsp. <i>flexuosa</i>	SG2341	107,570–114,787*	134,809–142,026*
<i>Lobelia ghiesbreghtii</i>	RG3015	107,337–114,284	134,265–141,206
<i>Lobelia goldmannii</i>	SAS235	107,982–115,508	135,536–143,062
<i>Lobelia gruina</i>	MJC30278	107,437–114,276	134,054–140,893
<i>Lobelia guerrerensis</i>	VWS5316	108,013–115,149	135,156–142,292
<i>Lobelia gypsophila</i>	TJA1752	109,513–116,862	136,572–143,915
<i>Lobelia hartwegii</i>	VWS5918	108,266–115,747	135,694–143,175
<i>Lobelia irasuensis</i> var. <i>fucata</i>	LLP350	134,855–142,300	107,556–115,001
<i>Lobelia jaliscensis</i>	TJA1776	107,587–114,984	134,935–142,332
<i>Lobelia knoblochii</i>	GR5574	108,532–115,644	135,665–142,777
<i>Lobelia laxiflora</i>	RGsn	107,836–114,975	134,990–142,129
<i>Lobelia laxiflora</i> var. <i>angustifolia</i>	TRVD658	108,347–115,678	135,693–143,024
<i>Lobelia lobelioides</i>	RCD2061	108,247–115,359	135,313–142,425
<i>Lobelia longicaulis</i>	JIC21345	107,384–114,316	134,396–141,328
<i>Lobelia macrocentron</i>	HAG324	108,081–115,154	135,185–142,258
<i>Lobelia margarita</i>	TJA1759	107,660–114,949	134,632–141,921
<i>Lobelia mcvaughii</i>	TJA348	108,597–115,748	135,778–142,929
<i>Lobelia nana</i>	CC72	107,025–113,906	133,648–140,529
<i>Lobelia plebeia</i>	EM19646	107,816–114,658	134,720–141,562
<i>Lobelia pulchella</i>	RGsn	107,323–114,363	134,402–141,442
<i>Lobelia sartorii</i>	RG3006	106,224–113,252	133,026–140,054
<i>Lobelia tatea</i>	RG3003	107,060–114,376	134,265–141,581
<i>Lobelia volcanica</i>	VWS4737	107,623–114,891	134,895–142,163
<i>Lobelia xalapensis</i>	TJA1790	107,253–114,239	133,606–140,592

Taxon	Voucher	ycf1 in IRB (bp)	ycf1 in IRA (bp)
<i>Lobelia yucatana</i>	GAM4737	107,145–114,464	129,373–136,692

infA gene variation

Variation in the plastid-encoded *infA* gene was detected among the 41 plastomes examined. *infA* was annotated as a CDS in 30 taxa, whereas five taxa (*L. anatina*, *L. ehrenbergii* subsp. *ehrenbergii*, *L. laxiflora*, *L. modesta*, and *L. nana*) contained a pseudogenized copy characterized by internal stop codons within the open reading frame. *infA* was absent from the plastome annotations of six taxa (*D. virgata*, *L. bryophila* var. *fimbriata*, *L. gruina*, *L. longicaulis*, *L. plebeia*, and *L. pulchella*). Four of the five taxa comprising clade C3 lacked *infA*, whereas additional absences occurred in C2 (*D. virgata*) and C11 (*L. gruina*). (Table 3).

Table 3
infA status in plastomes lacking an intact copy.

<i>infA</i> status	Taxon	Voucher
Absent	<i>Diastatea virgata</i>	TJA1772
	<i>Lobelia bryophila</i> var. <i>fimbriata</i>	VWS5660
	<i>Lobelia gruina</i>	MJC30278
	<i>Lobelia longicaulis</i>	JIC21345
	<i>Lobelia plebeia</i>	EM19646
	<i>Lobelia pulchella</i>	RGsn
Pseudogene	<i>Lobelia anatina</i>	TJA1622
	<i>Lobelia ehrenbergii</i> subsp. <i>ehrenbergii</i>	TJA1753
	<i>Lobelia laxiflora</i>	RGsn
	<i>Lobelia modesta</i>	TJA1350
	<i>Lobelia nana</i>	CC72

Simple sequence repeats

A total of 1,396 simple sequence repeats (SSRs) were identified across the sampled plastomes. Mononucleotide repeats were by far the most abundant class, accounting for approximately 79% of all SSRs detected (Fig. 4). Dinucleotide, trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide repeats occurred at substantially lower frequencies, with hexanucleotide repeats representing the least abundant class.

Phylogeny

The whole-plastome maximum-likelihood analysis was used as the primary phylogenetic reconstruction (Fig. 5), with the shared protein-coding gene (CDS) analysis providing a comparison based on coding regions (Figure S2). The two analyses yielded largely congruent topologies, with 11 clades (C1–C11) recognized in both. Four of the five *Lobelia* sections represented in the Mexican Transition Zone were recovered as monophyletic: sects. *Cryptostemon*, *Homochilus*, *Lobelia*, and *Tylomium*. In contrast, sect. *Stenotium* was polyphyletic, with its species distributed among several distinct clades. The three sampled species of *Diastatea* were not recovered as monophyletic but were distributed between two lineages (C1 and C2), each receiving 100/100 SH-aLRT/UFBoot support in both analyses. *Lobelia alanae* was not recovered with *L. cardinalis* and *L. inflata*, the sampled representatives of sect. *Lobelia*, in either analysis.

Plastome rearrangements and the absence of *infA* were distributed among different lineages of the phylogeny. The two major rearrangements identified in the comparative analyses occurred in species assigned to sect. *Stenotium*. The IR reduction was restricted to *L. yucatana* (C4), whereas the LSC inversion was restricted to *L. gypsophila* (C5). *infA* was absent from taxa distributed across three clades: *Diastatea virgata* (C2); *L. bryophila* var. *fimbriata*, *L. longicaulis*, *L. plebeia*, and *L. pulchella* (C3); and *L. gruina* (C11). Thus, the two major plastome rearrangements occurred in different clades and were each restricted to a single species, whereas the *infA* absence was distributed across multiple lineages.

Discussion

This study presents the first comparative plastome analysis of *Lobelia* from the Mexican Transition Zone (MTZ), including representatives of all five sections occurring in the region. Previous plastome studies in Lobelioideae have largely emphasized phylogenetic reconstruction rather than comparative plastome evolution. Although plastome-scale datasets have been used to infer relationships among the major lineages of Lobelioideae, including representatives of African giant lobelias, Hawaiian lobelioids, Southern Hemisphere lineages, and a limited number of Mexican taxa [18], comparative analyses of plastome structural evolution across the species-rich *Lobelia* assemblage of the MTZ have remained lacking. By substantially expanding taxonomic sampling across these five sections, our study provides the first comparative assessment of plastome structural variation and evolution within this assemblage.

Our results indicate that plastome structural evolution in *Lobelia* and *Diastatea* of the MTZ has been relatively conservative compared with the extensive structural variation documented more broadly across Campanulaceae. Previous studies have revealed substantial plastome rearrangements across major lineages of the family [17, 18], demonstrating that structural rearrangements have occurred repeatedly during plastome evolution in Campanulaceae.

In contrast, despite broad sampling across the five *Lobelia* sections occurring in the MTZ, we detected relatively limited structural variation, characterized primarily by a single large inversion and variation in IR length. This apparent contrast may partly reflect differences in phylogenetic scale and taxonomic breadth, as previous studies have examined plastome structural evolution across substantially deeper divisions within Campanulaceae and Lobelioideae [17, 18], whereas our sampling is concentrated within a regional assemblage of *Lobelia* and *Diastatea*. At this narrower evolutionary scale, plastome organization appears comparatively conserved, resembling the general pattern of structural conservation reported for many other angiosperm lineages [60].

Against this generally conserved background, the large LSC inversion in *L. gypsophila* is particularly informative. Its restriction to a single species indicates that substantial structural changes can arise within otherwise highly collinear plastomes and need not characterize entire sections or major clades. This pattern is consistent with the broader history of Lobelioideae, in which inversions have arisen repeatedly in different lineages [18, 36]. Thus, the *L. gypsophila* inversion provides an additional example of species-specific structural variation within Lobelioideae, occurring against a background of comparatively conserved plastome organization among the MTZ taxa examined here.

The genomic position of the *L. gypsophila* inversion also provides a potentially informative connection with previously described rearrangements in Lobelioideae. Knox et al. [36] found that nine of ten inversion endpoints examined occurred in intergenic spacers, seven adjacent to tRNA genes, including an endpoint in the *rpoB-trnC* intergenic region. This region has also been implicated in subsequent analyses of plastome rearrangements in Lobeliaceae [19]. The *L. gypsophila* inversion similarly involves the region surrounding *trnC-GCA*, indicating that this genomic region has previously been associated with plastome rearrangements. Rather than demonstrating a common mechanism, this correspondence raises the possibility that some tRNA-associated intergenic regions are recurrently involved in plastome structural change within Lobelioideae. Different structural rearrangements have also been reported in other *Lobelia* species, including the East Asian *L. zeylanica* and *L. heyneana* [25], further indicating that such changes have arisen independently in different parts of the genus. Large inversions are not restricted to Lobelioideae, however, and have been reported across angiosperms, including an approximately 25-kb double inversion in *Parthenium hysterophorus* (Asteraceae) [61].

Inverted repeat expansion and contraction represent another important source of plastome structural and size variation and, in some cases, can influence gene content [25]. Despite the extensive variation in IR boundaries documented within Campanulaceae, the LSC/IR and SSC/IR junctions were highly conserved across most plastomes examined here. The position of these junctions also reflects the distinctive history of IR evolution in Campanulaceae. In many angiosperms, *ndhF* occurs near the IRb/SSC junction, whereas in the plastomes examined here this boundary was associated with *ndhE*. This configuration is consistent with the previously documented expansion of the IR into the SSC in Campanulaceae, which extended the IR beyond *ndhF* and altered the positions of genes relative to the IR/SSC boundary [17, 22, 23]. *Lobelia yucatana* represented the principal exception, with substantially shorter IRs (32,782 bp), accompanied by the longest SSC region (12,230 bp) and the smallest plastome (160,815 bp) in our dataset. Together, these features indicate a distinct IR–SSC boundary configuration in *L. yucatana*, characterized primarily by a reduction in IR length and a corresponding expansion of the SSC region. This reduction in IR length also contributes to the smaller overall plastome size of this species.

Because *ycf1* commonly spans or occurs near the IR/SSC junction and is frequently represented by a complete copy together with a truncated copy associated with the opposite IR boundary, shifts in this region can also contribute to variation in *ycf1* length and copy structure [62]. The variation observed in *ycf1* among the MTZ plastomes therefore occurs in a genomic region closely associated with IR-boundary evolution. Importantly, shifts in IR boundaries may have evolutionary consequences beyond changes in plastome size and organization. Zhu et al. [63] showed that genes incorporated into the IR can experience reduced substitution rates, consistent with the enhanced copy-correction mechanisms associated with the duplicated IR [63]. Thus, changes in IR boundaries can potentially alter the evolutionary rates of genes transferred between single-copy and repeat regions. The lack of comparable boundary shifts among the

remaining taxa further indicates that pronounced IR variation is lineage-specific within the MTZ assemblage. Similar conservation of plastome organization and IR boundaries has been reported in several other angiosperm groups [64–67], illustrating that extensive structural rearrangement is not a universal feature of plastome evolution, even in groups with a broader history of structural change.

Variation in gene content was also evident in the plastid-encoded *infA* gene, which encodes translation initiation factor 1 and is among the protein-coding genes most frequently lost from angiosperm plastomes. Millen et al. [68] demonstrated that *infA* has become nonfunctional independently in numerous angiosperm lineages and documented multiple independent transfers of functional copies to the nuclear genome, indicating an unusually dynamic evolutionary history for this gene. Among the 41 newly assembled plastomes examined here, *infA* was annotated as a CDS in 30 taxa, occurred as a pseudogenized copy characterized by internal stop codons within the open reading frame in five taxa, and was absent in six. The phylogenetic distribution of these states suggests that changes in *infA* integrity and retention have occurred in multiple lineages rather than being restricted to a single clade. Four of the six taxa lacking *infA* were grouped in sect. *Stenotium* C3 (*L. bryophila* var. *fimbriata*, *L. longicaulis*, *L. plebeia*, and *L. pulchella*), whereas additional absences occurred in *Diastatea virgata* (C2) and *L. gruina* (C11). Among the three sampled species of *Diastatea*, *infA* was annotated as a CDS in *D. maximiliana* and *D. micrantha* but was absent in *D. virgata*, indicating that absence of the gene does not characterize *Diastatea* as a whole. The distribution of CDS, pseudogenized, and absent copies in our dataset therefore parallels the repeated loss of plastid *infA* documented across angiosperms [68], suggesting that the variation observed in *Lobelia* and *Diastatea* represents part of a broader pattern of evolutionary lability of this gene and may also characterize particular lineages within the MTZ assemblage.

A separate component of plastome variation examined in this study was the distribution of simple sequence repeats (SSRs), which are of particular interest because of their potential application as molecular markers. Mononucleotide repeats were by far the most abundant SSR class, accounting for approximately 79% of the 1,396 SSRs identified, whereas hexanucleotide repeats were the least frequent. The predominance of mononucleotide SSRs is consistent with patterns reported in other plastome studies [69]. Although their polymorphism was not evaluated here, the SSRs identified in these plastomes provide a set of candidate loci that could be further assessed as molecular markers for studies of genetic variation, population structure, and relationships among closely related *Lobelia* species.

The plastome phylogenetic analyses also provide important insight into relationships among the *Lobelia* lineages represented in the MTZ. Four of the five sampled sections (sects. *Cryptostemon*, *Homochilus*, *Lobelia*, and *Tylomium*) were recovered as monophyletic, providing support for these elements of the morphological infrageneric classification proposed by Lammers [30]. This resolution contrasts with previous phylogenetic studies based on more limited plastid-marker datasets or taxon sampling, in which relationships involving some of these groups received weak support or remained unresolved [26, 40]. The major sectional relationships recovered from the plastome data are also largely congruent with a recent nuclear phylogenomic analysis of *Lobelia* from the MTZ [35], which recovered the same four sections as monophyletic and sect. *Stenotium* as polyphyletic. Notably, *L. alanae*, which has been hypothesized to belong to sect. *Lobelia* based on seed coat sculpturing [70], was not recovered with that section in either genomic dataset. In the plastome phylogeny, *L. alanae* was instead recovered as sister to a strongly supported clade comprising representatives of sects. *Tylomium*, *Cryptostemon*, *Homochilus*, and one lineage of the polyphyletic sect. *Stenotium* (C11). Its placement outside sect. *Lobelia* in both the plastome phylogeny and the nuclear phylogeny [35] indicates that the infrageneric placement of *L. alanae* warrants further evaluation.

Comparison of the whole-plastome phylogeny with the nuclear phylogeny [35] revealed differences in the placement of some major lineages. For example, although sect. *Tylomium* was recovered as monophyletic in both phylogenies, its phylogenetic position differed between the plastid and nuclear topologies. Despite this topological difference, both phylogenies recovered sects. *Cryptostemon*, *Homochilus*, *Lobelia*, and *Tylomium* as monophyletic and sect. *Stenotium* as polyphyletic. The polyphyly of sect. *Stenotium* has also been recovered in previous phylogenetic studies using plastid markers and representatives from outside the MTZ [26, 27, 71], further supporting the non-monophyly of the section under its current circumscription.

Finally, the non-monophyly of *Diastatea* and its placement within a clade that also includes species currently assigned to sect. *Stenotium* (C3) indicate that the circumscription of these groups and their taxonomic relationships warrant further evaluation. More broadly, these relationships highlight the need to reassess generic and infrageneric boundaries within Lobelioideae using expanded taxon sampling and evidence from multiple genomic compartments. Together, these findings provide a plastome-based framework for understanding genome evolution and phylogenetic relationships among *Lobelia* and its allies in the MTZ, while offering an additional genomic perspective for future evaluation of taxonomic boundaries within Lobelioideae.

Conclusions

This study represents the first comparative plastome analysis of *Lobelia* from the Mexican Transition Zone (MTZ) to include representatives of all five sections occurring in the region. Using newly assembled and annotated plastomes, we found that genome structure, gene

content, gene order, and overall organization are highly conserved across the sampled taxa. Against this conserved genomic background, the principal structural differences were a large LSC inversion in *L. gypsophila* and a pronounced reduction of the IR in *L. yucatana*, both belonging to sect. *Stenotium*. Additional variation was observed in *infA*, which was annotated as a CDS in 30 taxa, occurred as a pseudogenized copy in five, and was absent in six. Plastome phylogenetic analyses recovered sects. *Homochilus*, *Cryptostemon*, *Lobelia*, and *Tylomium* as monophyletic, whereas sect. *Stenotium* was polyphyletic. The non-monophyly of *Diastatea* and its placement with species currently assigned to sect. *Stenotium* further highlight the need to evaluate generic and infrageneric boundaries within Lobelioideae. Overall, our results indicate that plastome evolution among *Lobelia* and *Diastatea* of the MTZ has been comparatively conservative across the phylogenetically diverse lineages sampled. These newly generated plastome resources and the resulting phylogenetic framework provide a foundation for future studies of genome evolution, systematics, and biogeography in Lobelioideae and Campanulaceae.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Conflict of interest

None

Funding

This work was supported by the National Science Foundation (NSF2242589) TJA and LPL.

Author Contribution

MAPP: Conceptualization; Methodology; Formal analysis; Field collections; Writing – original draft; Writing – review & editing. KGK: Methodology; Formal analysis; Writing – review & editing. NCK: Writing – review & editing. ACC: Field collections; Writing – review & editing. RGJH: Writing – review & editing. LPL: Funding acquisition; Conceptualization; Methodology; Writing – review & editing. TJA: Funding acquisition; Conceptualization; Methodology; Field collections; Supervision; Writing – review & editing.

Acknowledgement

We are grateful to the herbaria listed in methods sections who granted access to leaf tissue used in this study. We thank Wen-Bin Yu for helpful discussions on GetOrganelle outputs and GFA graphs. We also thank Advanced Research Computing and resources from the High Performance Computing Cluster at Northern Arizona University.

Data Availability

All the newly annotated plastomes have been deposited in GenBank.

References

1. Charboneau JLM, Cronn RC, Liston A, Wojciechowski MF, Sanderson MJ. Plastome Structural Evolution and Homoplastic Inversions in Neo-Astragalus (Fabaceae). *Genome Biol Evol* 2021, 13(10).
2. Yang T, Sahu SK, Yang L, Liu Y, Mu W, Liu X, Strube ML, Liu H, Zhong B. Comparative Analyses of 3,654 Plastid Genomes Unravel Insights Into Evolutionary Dynamics and Phylogenetic Discordance of Green Plants. *Front Plant Sci.* 2022;13:808156.
3. Birky CW Jr. Uniparental inheritance of mitochondrial and chloroplast genes mechanisms and evolution. *Proceedings of the National Academy of Sciences* 1995, 92:11331–11338.

4. Neale DB, Marshall KA, Sederoff RR. Chloroplast and mitochondrial DNA are paternally inherited in *Sequoia sempervirens* D. Don Endl. PNAS. 1989;86:9347–9.
5. Loeuille B, Thode V, Siniscalchi C, Andrade S, Rossi M, Pirani JR. Extremely low nucleotide diversity among thirty-six new chloroplast genome sequences from *Aldama* (Heliantheae, Asteraceae) and comparative chloroplast genomics analyses with closely related genera. PeerJ. 2021;9:e10886.
6. Gitzendanner MA, Soltis PS, Yi T-S, Li D-Z, Soltis DE. Plastome Phylogenetics: 30 Years of Inferences Into Plant Evolution. In: *Plastid Genome Evol* 2018: 293–313.
7. Ji Y, Yang L, Chase MW, Liu C, Yang Z, Yang J, Yang JB, Yi TS. Plastome phylogenomics, biogeography, and clade diversification of *Paris* (Melanthiaceae). BMC Plant Biol. 2019;19(1):543.
8. Kim YK, Jo S, Cheon SH, Kwak M, Kim YD, Kim KJ. Plastome evolution and phylogeny of subtribe *Aeridinae* (Vandaeae, Orchidaceae). Mol Phylogenet Evol. 2020;144:106721.
9. dePamphilis CW, Palmer JD. Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature. 1990;348:337–9.
10. Mukhopadhyay J, Hausner G. Organellar Introns in Fungi, Algae, and Plants. Cells 2021, 10(8).
11. Kim K-J, Lee H-L. Widespread Occurrence of Small Inversions in the Chloroplast Genomes of Land Plants. Mol Cells. 2005;19(1):104–13.
12. Lee C, Choi IS, Cardoso D, de Lima HC, de Queiroz LP, Wojciechowski MF, Jansen RK, Ruhlman TA. The chicken or the egg? Plastome evolution and an independent loss of the inverted repeat in papilionoid legumes. Plant J. 2021;107(3):861–75.
13. Cauz-Santos LA, da Costa ZP, Callot C, Cauet S, Zucchi MI, Berges H, van den Berg C, Vieira MLC. A Repertory of Rearrangements and the Loss of an Inverted Repeat Region in *Passiflora* Chloroplast Genomes. Genome Biol Evol. 2020;12(10):1841–57.
14. Katayama H, Ogiwara Y. Phylogenetic affinities of the grasses to other monocots as revealed by molecular analysis of chloroplast DNA. Curr Genet. 1996;29(6):572–81.
15. Jansen RK, Palmer JD. A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). Proc Natl Acad Sci. 1987;84:5818–22.
16. Wang M, Wang X, Sun J, Wang Y, Ge Y, Dong W, Yuan Q, Huang L. Phylogenomic and evolutionary dynamics of inverted repeats across *Angelica* plastomes. BMC Plant Biol. 2021;21(1):26.
17. Cosner ME, Raubeson LA, Jansen RK. Chloroplast DNA rearrangements in *Campanulaceae*: phylogenetic utility of highly rearranged genomes. BMC Evol Biol. 2004;4:27.
18. Knox EB. The dynamic history of plastid genomes in the *Campanulaceae* sensu lato is unique among angiosperms. Proc Natl Acad Sci USA. 2014;111(30):11097–102.
19. Knox EB. DNA Sequence Analysis of an Inversion Hot Spot in *Lobeliaceae* Plastomes. Plants (Basel) 2022, 11(21).
20. Benítez-Villaseñor A, Jost M, Granados Mendoza C, Wanke S, Meza-Lázaro RN, Peñafiel Cevallos M, Freire E, Magallón S. Exploring Structural Plastome Evolution in *Asterales*: Insights from Off-Target Hybrid Enrichment Data on the Small Single-Copy Region. J Mol Evol. 2025;93(1):111–23.
21. Hong CP, Park J, Lee Y, Lee M, Park SG, Uhm Y, Lee J, Kim CK. accD nuclear transfer of *Platycodon grandiflorum* and the plastid of early *Campanulaceae*. BMC Genomics. 2017;18(1):607.
22. Haberle RC, Fourcade HM, Boore JL, Jansen RK. Extensive rearrangements in the chloroplast genome of *Trachelium caeruleum* are associated with repeats and tRNA genes. J Mol Evol. 2008;66(4):350–61.
23. Knox EB, Palmer JD. The chloroplast genome arrangement of *Lobelia thuliniana* (Lobeliaceae) expansion of the inverted repeat in an ancestor of the *Campanulales*. Plant Syst Evol. 1999;214:49–64.
24. Lopez-Caamal A, Gandee T, Galloway LF, Barnard-Kubow KB. Substantial structural variation and repetitive DNA content contribute to intraspecific plastid genome evolution. BMC Genomics. 2025;26(1):340.
25. Li CJ, Xie XT, Liu HX, Wang RN, Li DZ. Plastome evolution in the East Asian *lobelias* (Lobelioideae) using phylogenomic and comparative analyses. Front Plant Sci. 2023;14:1144406.
26. Kagame SP, Gichira AW, Chen LY, Wang QF. Systematics of *Lobelioideae* (Campanulaceae): review, phylogenetic and biogeographic analyses. PhytoKeys. 2021;174:13–45.
27. Antonelli A. Higher level phylogeny and evolutionary trends in *Campanulaceae* subfam. *Lobelioideae*: molecular signal overshadows morphology. Mol Phylogenet Evol. 2008;46(1):1–18.
28. Knox EB, Muasya AM, Muchhala N. The predominantly South American clade of *Lobeliaceae*. Syst Bot. 2008;33(2):462–8.

29. Crowl AA, Miles NW, Visger CJ, Hansen K, Ayers T, Haberle R, Cellinese N. A global perspective on Campanulaceae: Biogeographic, genomic, and floral evolution. *Am J Bot.* 2016;103(2):233–45.
30. Lammers TG. Revision of the Infrageneric Classification of *Lobelia* L. (Campanulaceae: Lobelioideae). *Ann Mo Bot Gard.* 2011;98(1):37–62.
31. Rzedowski J. Updated inventory of the Mexican species of Campanulaceae. *Bot Sci.* 2019;97(1):110–27.
32. Morrone JJ. Fundamental biogeographic patterns across the Mexican Transition Zone: an evolutionary approach. *Ecography.* 2010;33(2):355–61.
33. Althaus KN, González-Elizondo MS, González-Rodríguez A, Rodríguez Correa H, Hipp AL. Phylogenetic Niche Conservatism Drives Floristic Assembly Across Mexico's Temperate-Tropical Divide. *Glob Ecol Biogeogr* 2025, 34(4).
34. Villaseñor JL, Ortiz E, Delgadillo-Moya C, Juárez D. The breadth of the Mexican Transition Zone as defined by its flowering plant generic flora. *PLoS ONE.* 2020;15(6):e0235267.
35. Pérez-Pérez MA, Crowl AA, Ball L, Castro-Castro A, Hodel RG, Lagomarsino LP, Ayers TJ. Incomplete lineage sorting shapes phylogenomic discordance in Mexican *Lobelia* (Campanulaceae). *Available at SSRN 7069035* 2026.
36. Knox EB, Downie SR, Palmer JD. Chloroplast genome rearrangements and the evolution of giant lobelias from herbaceous ancestors. *Mol Biol Evol.* 1993;10:414–30.
37. Givnish TJ. Giant lobelias exemplify convergent evolution. *BMC Biol.* 2010;8:3.
38. Antonelli A. Have giant lobelias evolved several times independently? Life form shifts and historical biogeography of the cosmopolitan and highly diverse subfamily Lobelioideae (Campanulaceae). *BMC Biol.* 2009;7:82.
39. Johnson EP, Ayers TJ. Molecular Phylogeny and Taxonomic Revision of the Genus *Diastatea* (Campanulaceae: Lobelioideae). *Syst Bot.* 2022;47(1):61–84.
40. Koopman MM, Ayers T. Nectar spur evolution in the Mexican Lobelias. *Am J Bot.* 2005;92(3):558–62.
41. Uribe-Convers S, Carlsen MM, Lagomarsino LP, Muchhala N. Phylogenetic relationships of *Burmeistera* (Campanulaceae: Lobelioideae): Combining whole plastome with targeted loci data in a recent radiation. *Mol Phylogenet Evol.* 2017;107:551–63.
42. Knox EB, Li C. The East Asian origin of the giant lobelias. *Am J Bot.* 2017;104(6):924–38.
43. Pérez-Pérez MA, Yu WB. Pleistocene origin and colonization history of *Lobelia columnaris* Hook. f. (Campanulaceae: Lobelioideae) across sky islands of West Central Africa. *Ecol Evol.* 2021;11(22):15860–73.
44. Inglis PW, Pappas MCR, Resende LV, Grattapaglia D. Fast and inexpensive protocols for consistent extraction of high quality DNA and RNA from challenging plant and fungal samples for high-throughput SNP genotyping and sequencing applications. *PLoS ONE.* 2018;13(10):e0206085.
45. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34(17):i884–90.
46. Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol* 2020, 21(1).
47. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics.* 2015;31(20):3350–2.
48. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* 2017;45(W1):W6–11.
49. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* 2019;47(W1):W59–64.
50. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, 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.
51. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 2004;14(7):1394–403.
52. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754–60.
53. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. Genome Project Data Processing S: The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078–9.
54. Li H, Guo Q, Xu L, Gao H, Liu L, Zhou X. CPJSdraw: analysis and visualization of junction sites of chloroplast genomes. *PeerJ.* 2023;11:e15326.
55. Huang L, Yu H, Wang Z, Xu W. CPStools: A package for analyzing chloroplast genome sequences. *iMetaOmics* 2024.

56. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30(4):772–80.
57. Borowiec ML. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ.* 2016;4:e1660.
58. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009;25(15):1972–3.
59. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol.* 2020;37(5):1530–4.
60. Palmer JD, Stein DB. Conservation of chloroplast genome structure among vascular plants. *Curr Genet.* 1986;10(11):823–33.
61. Lubna AS, Jan R, Asif S, Bilal S, Khan AL, Al-Rawahi AN, Kim KM, Al-Harrasi A. The complete plastome sequences of invasive weed *Parthenium hysterophorus*: genome organization, evolutionary significance, structural features, and comparative analysis. *Sci Rep.* 2024;14(1):4006.
62. Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, Cheng T, Guo J, Zhou S. ycf1, the most promising plastid DNA barcode of land plants. *Sci Rep.* 2015;5:8348.
63. Zhu A, Guo W, Gupta S, Fan W, Mower JP. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. *New Phytol.* 2016;209(4):1747–56.
64. Abdullah RA, Chen C, Yan R, Naz I, Heidari P, Ahmed I, Vallada A, Tian X. Comparative chloroplast genomics of *Erigeron* (Asteroideae, Asteraceae). *BMC Plant Biol.* 2025;25(1):1356.
65. Abdullah MF, Rahim A, Heidari P, Ahmed I, Poczar P. Comparative plastome analysis of *Blumea*, with implications for genome evolution and phylogeny of Asteroideae. *Ecol Evol.* 2021;11(12):7810–26.
66. Iram S, Hayat MQ, Tahir M, Gul A, Abdullah, Ahmed I. Chloroplast Genome Sequence of *Artemisia scoparia*: Comparative Analyses and Screening of Mutational Hotspots. *Plants (Basel)* 2019, 8(11).
67. Gichira AW, Avoga S, Li Z, Hu G, Wang Q, Chen J. Comparative genomics of 11 complete chloroplast genomes of Senecioneae (Asteraceae) species: DNA barcodes and phylogenetics. *Bot Stud.* 2019;60(1):17.
68. Millen RS, Olmstead RG, Adams KL, Palmer JD, Lao NT, Heggie L, Kavanagh TA, Hibberd JM, Gray JC, Morden CW. Many parallel losses of *infA* from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. *Plant Cell.* 2001;13(3):645–58.
69. Li PW, Lu YB, Qin XM, Zhang Q. Plastome phylogenomics unravels the evolutionary relationships and biogeographic history of Chloranthaceae. *BMC Plant Biol.* 2025;25(1):543.
70. Pérez-Pérez MA, Ayers TJ, Amith JD. A new species of *Lobelia* (Campanulaceae: Lobelioideae) from the Sierra Madre Oriental, Mexico. *Phytotaxa.* 2022;568(1):1–7.
71. Knox EB, Muasya AM, Phillipson PB. The Lobeliaceae originated in southern Africa. In: *Taxonomy and ecology of African plants, their conservation and sustainable use: 2006; Addis Ababa, Ethiopia*. Royal Botanic Gardens, Kew, U. K.: 215–27.

Figures

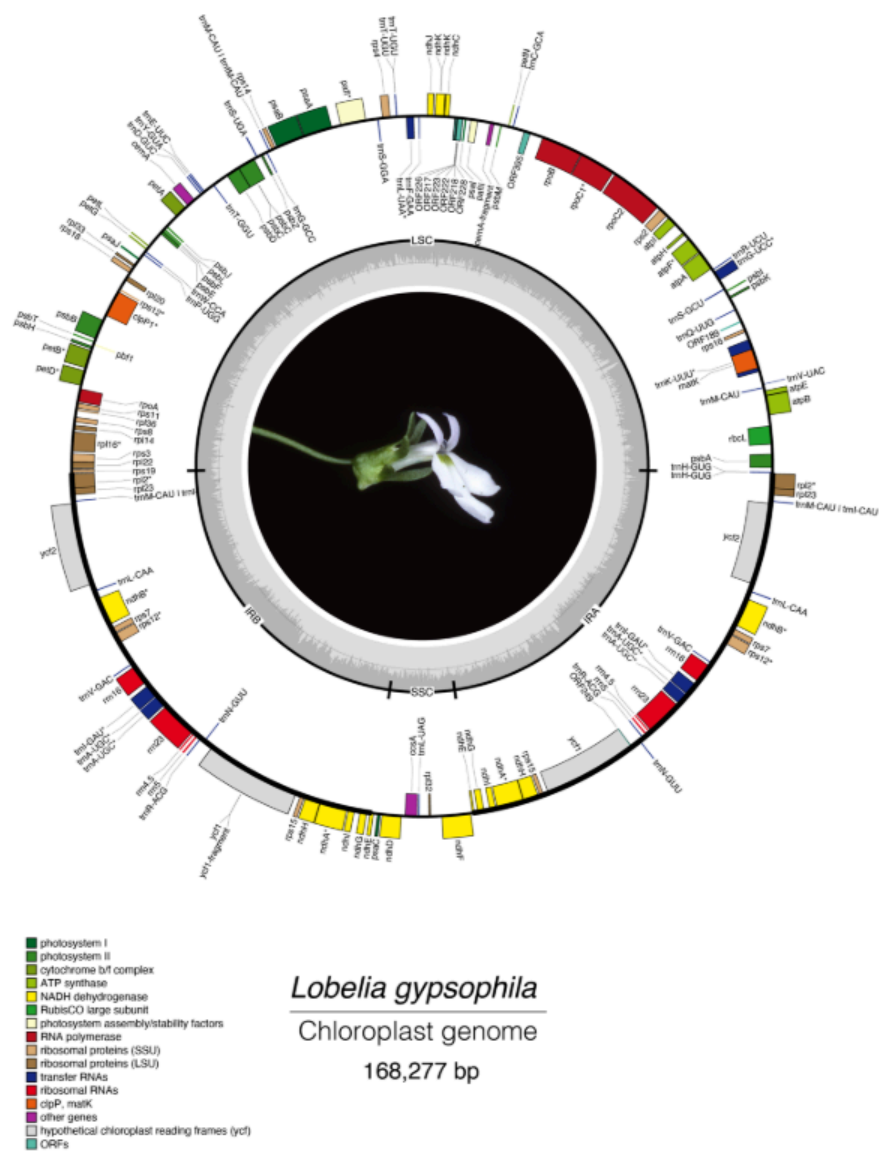


Figure 1

Annotated plastome of *Lobelia gypsophila* (sect. *Stenotium*), showing the large single-copy (LSC), small single-copy (SSC), and inverted repeat (IRA and IRB) regions.



Figure 2

Mauve alignment of seven representative plastomes of *Lobelia* sect. *Stenotium*, sect. *Tylomium*, *L. alanae*, and *Diastatea virgata*. The corresponding Mauve alignment including all plastomes is provided in Figure S3.

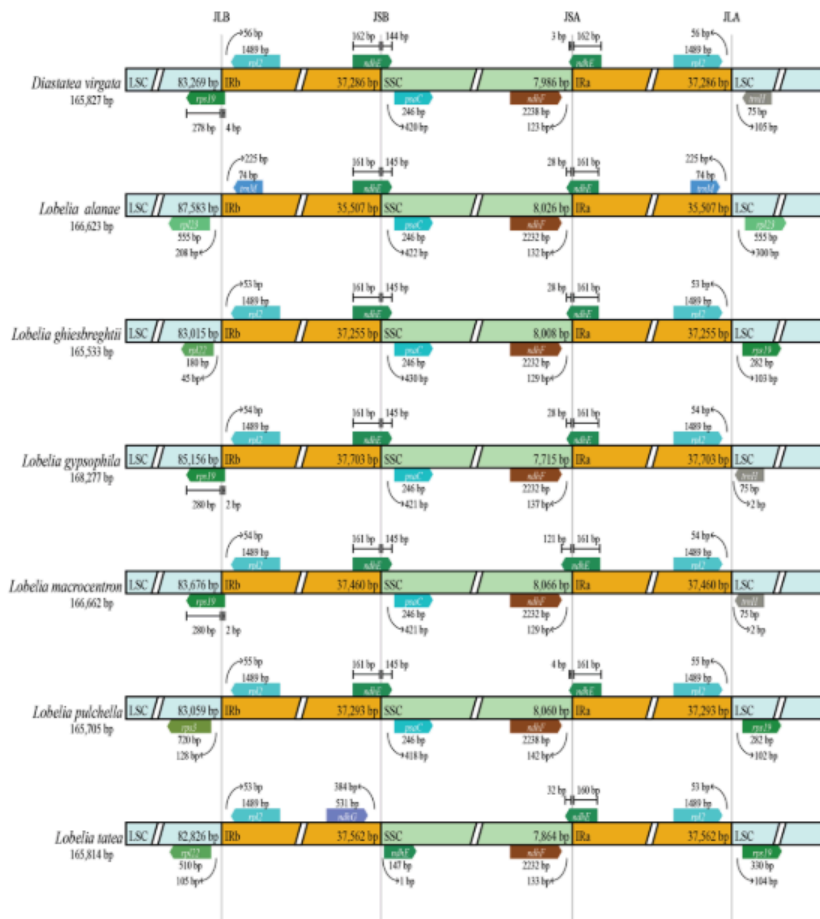


Figure 3

Inverted repeat (IR) boundaries among seven representative plastomes from the Mexican Transition Zone (MTZ). The corresponding IR boundary comparison including all plastomes is provided in Figure S4.

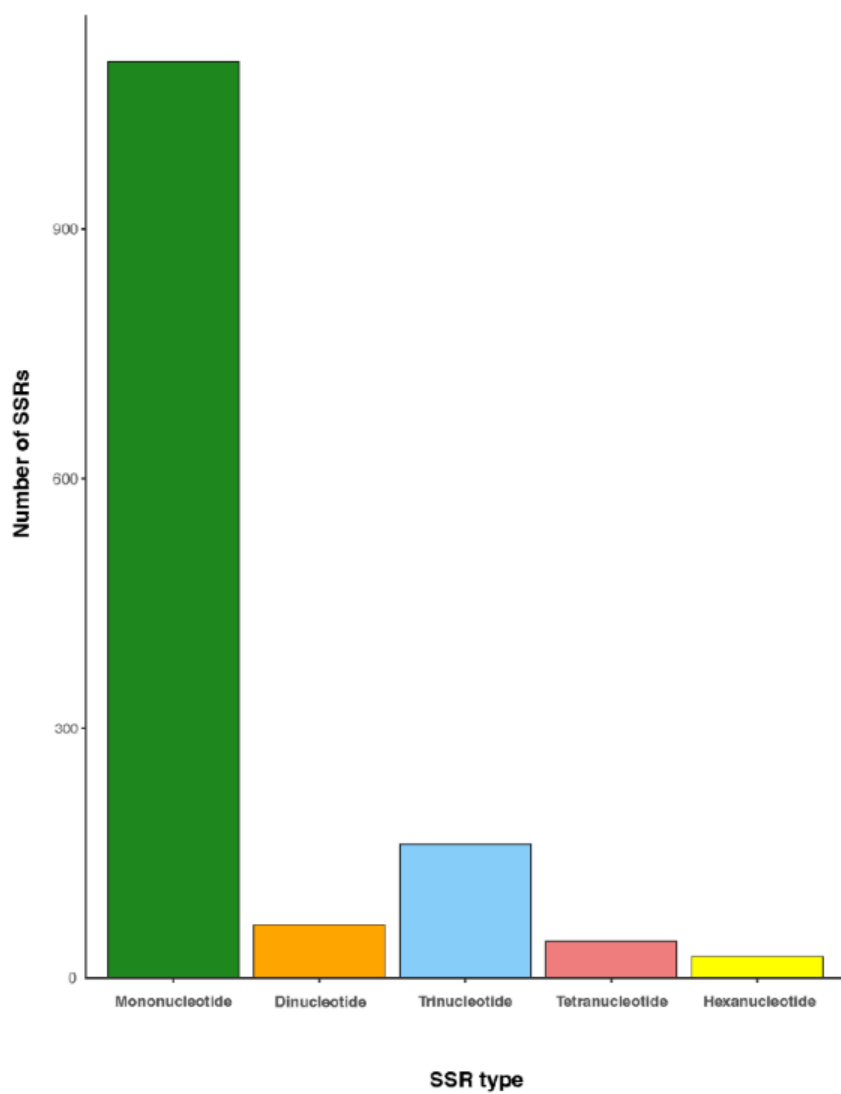


Figure 4

Composition of simple sequence repeat (SSR) motif types across 41 plastomes of *Lobelia* and *Diastatea*.

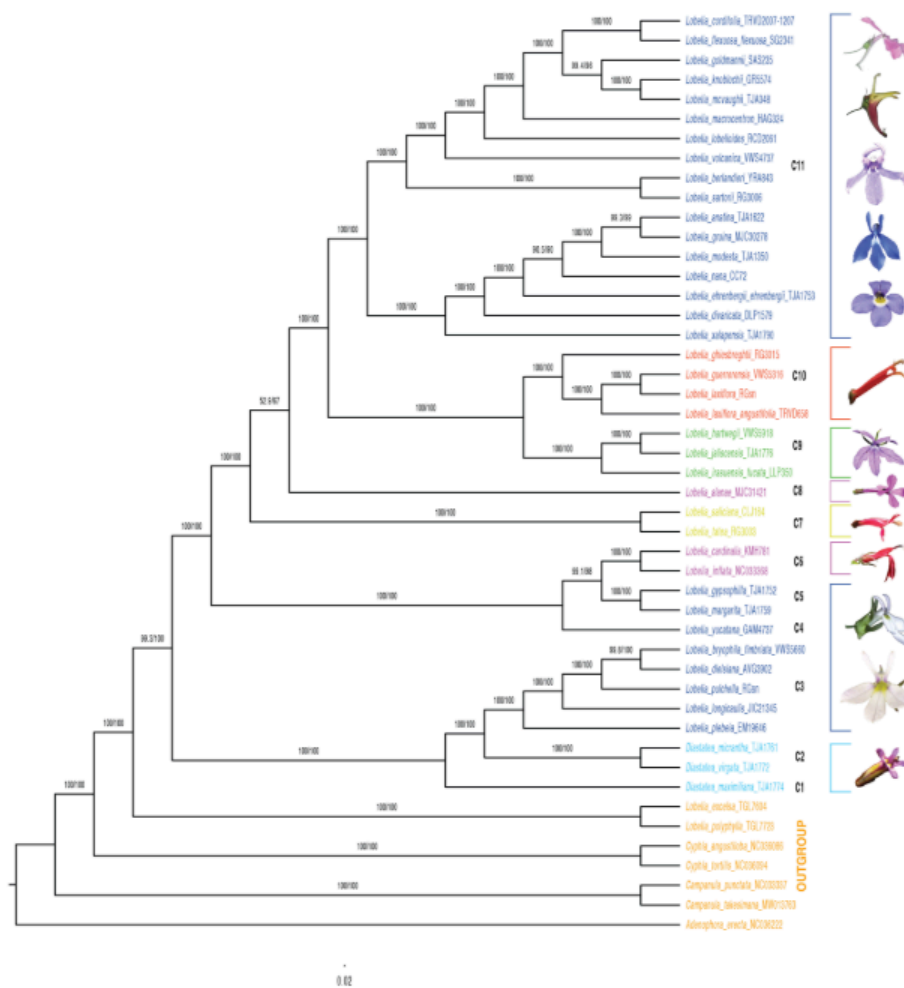


Figure 5

Maximum-likelihood phylogeny inferred from the concatenated whole-plastome alignment of *Lobelia* and related taxa from the Mexican Transition Zone (MTZ). Node values represent SH-aLRT support and ultrafast bootstrap support (SH-aLRT/UFBboot).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.ReadalignmentdepthinversioninLobeliagypsophila.png](#)
- [FigureS2.MaximumlikelihoodCDSconcatenatedofLobeliaandalliesoftheMTZ.png](#)
- [FigureS3.MauvealignmentwithplastomeofLobeliaoftheMTZ.png](#)
- [FigureS4.InvertedrepeatboundariesamongplastomesofLobeliaoftheMTZ.png](#)
- [TableS1.Voucherinformation.docx](#)