



OPEN ACCESS

EDITED BY

Zhiqiang Wu,
Chinese Academy of Agricultural
Sciences, China

REVIEWED BY

Junhu Kan,
Chinese Academy of Agricultural
Sciences, China
Shiyun Han,
Anhui Normal University, China

*CORRESPONDENCE

Mohamed Fokar
✉ m.fokar@ttu.edu
Younes Abbas
✉ a.younes@usms.ma

RECEIVED 11 January 2026

REVISED 24 February 2026

ACCEPTED 09 March 2026

PUBLISHED 01 April 2026

CITATION

Taha A, Rabeh K, Lamara M, Gaboun F,
Laads I, Lebki N, Diria G, Abdelwahd R,
Saghir K, Ettaqy A, El-Mderssa M,
Khamar H, Johnson MG, Fokar M and
Abbas Y (2026) Complete chloroplast
genome of *Euphorbia resinifera*:
overcoming biogeographical bias in
phylogenetic inference and establishing
a conservation genomics framework for
threatened North-West African
cactiform species.
Front. Plant Sci. 17:1785579.
doi: 10.3389/fpls.2026.1785579

COPYRIGHT

© 2026 Taha, Rabeh, Lamara, Gaboun,
Laads, Lebki, Diria, Abdelwahd, Saghir,
Ettaqy, El-Mderssa, Khamar, Johnson,
Fokar and Abbas. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance
with accepted academic practice. No
use, distribution or reproduction is
permitted which does not comply with
these terms.

Complete chloroplast genome of *Euphorbia resinifera*: overcoming biogeographical bias in phylogenetic inference and establishing a conservation genomics framework for threatened North-West African cactiform species

Abdelhakim Taha¹, Karim Rabeh², Mebarek Lamara³,
Fatima Gaboun⁴, Ilham Laads⁵, Nezha Lebki¹, Ghizlane Diria⁴,
Rabha Abdelwahd⁴, Karim Saghir¹, Abderrahim Ettaqy¹,
Mohamed ElMderssa¹, Hamid Khamar⁶, Matthew G. Johnson⁷,
Mohamed Fokar^{5*} and Younes Abbas^{1*}

¹Sultan Moulay Slimane University, Polydisciplinary Faculty of Beni Mellal, Applied Biology for Health,
Environment and Sustainable Development Laboratory, Beni Mellal, Morocco, ²Oasis System Research
Unit, Regional Center of Agricultural Research of Errachidia, National Institute of Agricultural Research,
Errachidia, Morocco, ³Institut de recherche sur les forêts, Université du Québec en Abitibi-
Témiscamingue, Rouyn-Noranda, QC, Canada, ⁴Biotechnology Unit, Regional Center of Agricultural
Research of Rabat, National Institute of Agricultural Research, Rabat, Morocco, ⁵Center for Biotechnology
and Genomics, Texas Tech University, Lubbock, TX, United States, ⁶Department of Botany and Plant
Ecology, Scientific Institute, Mohammed V University, Rabat, Morocco, ⁷Department of Biological
Sciences, Texas Tech University, Lubbock, TX, United States

Introduction: *Euphorbia resinifera* O. Berg (1863) is an endemic cactiform species
from Morocco's arid High Atlas Mountains, valued for its medicinal properties and
melliferous importance. It faces conservation challenges from habitat
degradation and phylogenetic incongruence between xeromorphic adaptations
and current sectional classification.

Methods: We pioneered a non-destructive spine-DNA genome-skimming
approach to generate the complete chloroplast genome, and applied a multi-
pronged molecular strategy including cpSSR profiling, whole-plastome
phylogenetics, and *rpl22* gene analysis.

Results: The complete chloroplast genome (163,065 bp; 35.11% GC) contains 132
genes and 96 SSR loci, representing the first genomic resource for threatened
North-West African cactiform *Euphorbia*. cpSSR profiles recovered relationships
only within sect. *Euphorbia*, while whole-plastome analysis showed
biogeographical clustering. The *rpl22* gene precisely distinguished growth
forms and revealed perfect sequence identity between morphologically similar
pairs (*E. ampliphylla* and *E. resinifera*, *E. neospinescens* and *E. neoarborescens*,
E. micractina and *E. pekinensis*), with *E. royleana* showing near-perfect identity to

E. ampliphylla and *E. resinifera* (1 SNP difference). Notably, *E. drupifera* was excluded from the cactiform core within sect. *Euphorbia* through 10 diagnostic SNPs and a unique AGC deletion.

Discussion: This study establishes *rpl22* as a powerful taxonomic discriminator for *Euphorbia* systematics, with future multi-gene approaches recommended for broader phylogenetic resolution.

KEYWORDS

Euphorbia resinifera, complete chloroplast genome, genome skimming, cactiform clade, *rpl22* gene, phylogenetic analysis, cpSSR analysis, plant systematics

1 Introduction

The genus *Euphorbia*, with over 2,000 species, constitutes the second largest angiosperm genus and forms a monophyletic group subdivided into four distinct subgenera (Wei et al., 2021). The subgenus *Euphorbia* comprises 21 sections, with sect. *Euphorbia* being the most morphologically variable, including over 340 species across Africa and South to Southeast Asia. Recent molecular phylogenetic studies using chloroplast markers (*matK*, *ndhF*, *rbcl*, *rpl16*, *trnL-F*) combined with nuclear ITS sequences have significantly advanced sectional classification within subgenus *Euphorbia* (Bruyns et al., 2006; Horn et al., 2012; Dorsey et al., 2013). However, species delimitation within sect. *Euphorbia* remains challenging due to conflicting morphological and molecular evidence, particularly where current phylogenetic markers fail to separate cactiform from pachycaul growth forms, as demonstrated by the problematic placement of pachycaul *E. drupifera* among morphologically distinct cactiform lineages (Horn et al., 2012; Dorsey et al., 2013).

Within sect. *Euphorbia*, cactiform xerophytes with ribbed stems represent a morphologically distinct group that resembles columnar Cactaceae but is distinguished by its latex secretion, cyathia inflorescences, and paired spines (Horn et al., 2012; Dorsey et al., 2013; Chaudhary et al., 2023). These species originated in Africa and occupy diverse habitats ranging from arid and semi-arid to humid tropical environments, most often on rocky or sandy soils (Horn et al., 2012, 2014; Ettaqy et al., 2020). North-west African species (*E. resinifera*, *E. officinarum*, *E. canariensis*, *E. handiensis*) share ecological and phylogenetic traits with East African and Arabian Peninsula taxa (Dorsey et al., 2013). Notably, *E. makallensis*, the East African cactiform most morphologically similar to North-West African *E. resinifera*, is of very limited distribution and declining in its native Ethiopian highlands (Wilson and Munro, 2019). Most xerophytic cactiform *Euphorbia* are declining in arid to hyperarid zones characterized by prolonged drought periods (Masrahi et al., 2015; Al-Qthanin and Alharbi, 2023; Taha et al., 2023).

Euphorbia resinifera O. Berg (1863), endemic to Morocco's High Atlas and Anti-Atlas Mountains, exemplifies these taxonomic uncertainties. This multipurpose species, valued for medicinal properties and apiculture, has undergone rapid decline over recent decades, driven by increasing aridity and direct human pressure, potentially leading to extinction (Taha et al., 2023, 2024). It is currently classified as "Near Threatened" in Morocco's

Red Book of Vascular Flora (Fennane, 2018). Traditional tissue sampling approaches are problematic for cactiform species due to their leafless morphology and irritant latex, with permission to sample tissue particularly difficult to obtain for rare and threatened species held in living collections. However, spine-derived DNA extraction offers a practical non-destructive alternative for genomic studies of threatened succulent species (Fehlberg et al., 2013). Genome-skimming approaches are particularly effective for organellar genome recovery (Straub et al., 2012) and well-suited for degraded DNA samples (Li et al., 2021).

The chloroplast genome is a powerful tool for plant phylogenomics due to its conserved quadripartite structure, moderate substitution rates, maternal inheritance, and rare recombination events (Ruhlman and Jansen, 2014; Zhao et al., 2024). Within slow-evolving plastomes, chloroplast simple sequence repeats (cpSSRs) provide particularly valuable taxonomic signatures. Furthermore, cpSSR distribution patterns including their abundance, density, and GC content exhibit lineage-specific signatures that reflect underlying taxonomic relationships (Zhu et al., 2021). Chloroplast SSRs have proven particularly effective for resolving species relationships and characterizing phylogenetic patterns within complex plant genera (Vieira et al., 2016). Recent advances in plastome-based systematics have demonstrated the effectiveness of integrating complete genome sequences with targeted marker analysis for resolving complex evolutionary relationships across diverse angiosperm families (Kan et al., 2024; Dong et al., 2025).

To address conservation needs and clarify the taxonomic position of these threatened species, we sequenced and characterized the complete chloroplast genome of *E. resinifera* using a genome-skimming approach applied to DNA extracted from plant spines, the first documented non-destructive plastome sequencing protocol for cactiform *Euphorbia* species. Specifically, our objectives were to: (1) generate and annotate the first complete chloroplast genome of a leafless shrub cactiform *Euphorbia* with ribbed stem from the arid zones of North-West Africa and Arabia; (2) examine plastome structural features and IR/SC boundary dynamics within section *Euphorbia*; (3) characterize SSR abundance and distribution patterns across the chloroplast genome; and (4) identify targeted plastid loci with high taxonomic resolution for distinguishing cactiform xerophyte lineages from pachycaul species. Together, these results provide a conservation-friendly systematic framework for resolving taxonomic relationships and supporting molecular monitoring of threatened cactiform *Euphorbia* lineages from arid and hyperarid regions.

2 Materials and methods

2.1 Plant material collection and genomic DNA extraction

Due to its high latex content, *Euphorbia resinifera* shows poor regenerative capacity and is undergoing a progressive decline. To avoid damaging spiny cactiforms *Euphorbia* species through stem cutting during sampling, a non-destructive sampling method was previously introduced by Fehlbeg et al. (2013) for this type of spiny cacti plant. Following this approach, samples consisting of 5 to 10 pairs of spines were collected from the upper part of the stems of *E. resinifera* individuals in the Fout-el-Anceur region near the city of Beni Mellal, Morocco (32°22.1140'N, 6°14.8390'W). The number of spine pairs collected per individual varied according to spine size to ensure sufficient material for DNA extraction. Spines were clipped using sterile eyebrow forceps and immediately placed into paper bags for transport.

The sampled population was considered homogeneous, consisting of semi-spherical bushes with tightly packed stems. A voucher specimen was deposited in the herbarium of the Scientific Institute, Mohammed V University, Rabat, Morocco (voucher number: RAB114831). The complete annotated chloroplast genome sequence has been deposited in NCBI GenBank under accession number PZ168867.

Genomic DNA was extracted from 5 to 10 spines, corresponding to approximately 20 mg of frozen and ground tissue, into a 2 mL microcentrifuge tube. DNA extraction was performed using the DNeasy Plant Mini Kit (Qiagen, Morrisville, USA) following the manufacturer's instructions. DNA concentration and purity were assessed using a Nanodrop spectrophotometer, and the extracted DNA was stored at -20 °C until sequencing.

2.2 DNA sequencing, chloroplast genome assembly, and annotation

DNA libraries were prepared from 50 ng of genomic DNA using the NEBNext Ultra II FS DNA Library Prep Kit (New England Biolabs). The resulting libraries were sequenced on an Illumina NovaSeq 6000 platform (S2 flow cell); to generate 2 × 150 bp paired-end reads. Raw sequencing reads were adapter-trimmed and quality-filtered using Trimmomatic v0.39 (Bolger et al., 2014). Read quality was then assessed using FastQC v0.12.1 (Andrews, 2010).

The chloroplast genome was *de novo* assembled using NOVOPlasty v4.3.5 (Dierckx et al., 2020), with the *rbcl* gene of *Euphorbia ampliphylla* used as the seed sequence. Assembly quality was validated using GetOrganelle v1.7.7.0 (Jin et al., 2020). Genome annotation was performed using GeSeq v2.03 (Tillich et al., 2017), CPGAVAS2 (Shi et al., 2019), and OGDRAW (Greiner et al., 2019).

2.3 SSR analysis and IR/SC boundary characterization

Simple sequence repeats (SSRs) were identified using MISA (Beier et al., 2017), with minimum repeats thresholds of 10, 6, 4, 4,

3, and 3 units for mono-, di-, tri-, tetra-, penta-, and hexanucleotides motifs, respectively. The boundaries of the inverted repeat (IR) regions and the junction between the large single-copy (LSC) and small single-copy (SSC) regions were comparatively analyzed for *E. resinifera* and 19 related *Euphorbia* species using IRscope (Díez Menéndez et al., 2023).

2.4 Phylogenetic and compositional analyses

2.4.1 SSR-based compositional analysis

Simple sequence repeat motif presence and absence were encoded as a binary matrix (1/0) for ten representative *Euphorbia* species (a subset of the 22 total spanning all four subgenera and major morphological forms). Sørensen dissimilarity indices were calculated and used for hierarchical clustering based on Ward's minimum variance method (Ward.D2) (Sørensen, 1948; Ward, 1963). Clustering robustness was assessed via Mantel tests (999 permutations) comparing SSR-based and phylogenetic distance matrices. Additionally, T/A ratios and A+T nucleotide composition were computed to quantify compositional divergence among four *Euphorbia* species within section *Euphorbia*.

2.4.2 Chloroplast genome-based phylogenetic analysis

Chloroplast genome sequences from 22 *Euphorbia* species (including *E. resinifera*) and three Euphorbiaceae outgroups (*Mallotus peltatus*, *Hevea brasiliensis*, and *Croton tiglium*) were aligned using MAFFT v7.505 (Katoh et al., 2018). Maximum likelihood phylogenies were inferred with IQ-TREE v2.0.7 (Minh et al., 2020) under the GTR+G substitution model with 1,000 bootstrap replicates to assess node support. Resulting trees were visualized using iTOL v6 (Letunic and Bork, 2024) and FigTree v1.4.5 (Rambaut et al., 2018). Genome-wide nucleotide diversity (Pi) was calculated across the 22 *Euphorbia* species using DnaSP v6.12.03 (Rozas et al., 2017) with a sliding window approach (window length = 600 bp, step size = 200 bp) to identify highly variable plastid regions. For *rpl22*-specific analyses, gene sequences were downloaded from NCBI and aligned with MAFFT, with phylogenetic trees constructed using IQ-TREE. Sequence variation, including single-nucleotide polymorphisms, indels, diagnostic trinucleotide motifs, and T/A ratios, was characterized via BLASTN (Altschul et al., 1990) to evaluate its taxonomic discrimination capacity within sect. *Euphorbia*.

3 Results

3.1 Structural features of the *E. resinifera* chloroplast genome

The chloroplast genome of *Euphorbia resinifera* was assembled from 13,167,418 sequencing reads, of which 6.97% (917,770 reads) corresponded to organelle DNA. A total of 904,376 reads were

TABLE 1 Characteristics of the complete chloroplast genome of *Euphorbia resinifera*.

Category	Items	Characteristics
Construction of chloroplast genome	LSC region (bp)	91,462
	SSC region (bp)	18,285
	IRa region (bp)	26,659
	IRb region (bp)	26,659
	Size of chloroplast genome (bp)	163,065
Gene content	Protein-coding genes	87
	tRNA	37
	rRNA	8
	Duplicate genes	19
	Genes on LSC region	82
	Genes on SSC region	12
	Genes on IRa region	19
	Genes on IRb region	19
	Total genes	132
GC content	GC content of LSC region (%)	32.06
	GC content of SSC region (%)	29.13
	GC content of IRa region (%)	42.39
	GC content of IRb region (%)	42.39
	Overall GC content (%)	35.11

successfully retained in the final assembly, resulting in an average coverage depth of 849×.

The complete plastome is 163,065 bp in length with an overall GC content of 35.11% and exhibits the typical quadripartite structure (Figure 1), consisting of a large single-copy (LSC) region of 91,462 bp (32.06% GC), a small single-copy (SSC) region of 18,285 bp (29.13% GC), and two inverted repeats (IRs) of 26,659 bp each (42.39% GC). The genome encodes a total of 132 annotated genes: 87 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. Gene distribution comprises 82 genes in the LSC region, 12 in the SSC region, and 19 genes in each IR region. Eight protein-coding genes are duplicated in the IR regions (*ndhB*, *rpl2*, *rpl23*, *rps7*, *rps12*, *rps19*, *ycf2*, *ycf15*), along with 7 tRNA genes and 4 rRNA genes. (Tables 1, 2).

3.2 Simple sequence repeats analysis

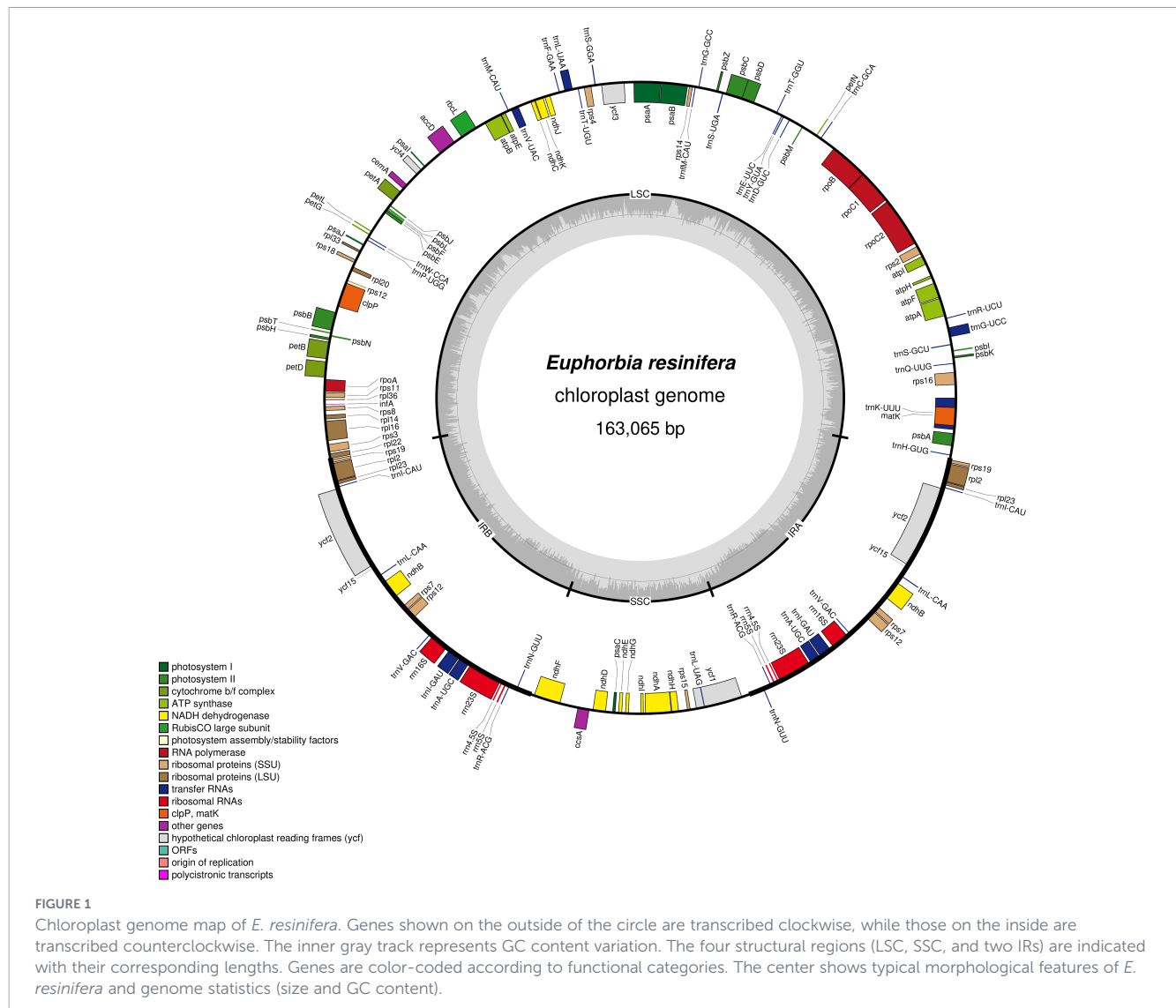
3.2.1 Analysis of SSR characteristics and distribution in the *E. resinifera* plastome

A comprehensive analysis of the *E. resinifera* chloroplast genome detected 96 SSRs with a high density of 0.59 SSRs/kb. Mononucleotide repeats predominated (83 SSRs, 86.5%), showing a

TABLE 2 Genes annotated in the chloroplast genome of *Euphorbia resinifera*.

Category of genes	Group of genes	Gene name
Photosynthesis-related Genes	Photosystem II	<i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i>
	Large subunit of rubisco	<i>rbcL</i>
	Photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>
	NADH dehydrogenase	<i>ndhA</i> *, <i>ndhB</i> *, <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
	Cytochrome b6/f complex	<i>petA</i> , <i>petB</i> , <i>petD</i> , <i>petG</i> , <i>petL</i> , <i>petN</i>
	ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> *, <i>atpH</i> , <i>atpI</i>
Transcription and translation related genes	Small subunit of ribosomal proteins	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> , <i>rps8</i> , <i>rps11</i> , <i>rps12</i> *, <i>rps14</i> , <i>rps15</i> , <i>rps18</i> , <i>rps19</i>
	Large subunit of ribosomal proteins	<i>rpl2</i> *, <i>rpl14</i> , <i>rpl16</i> , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> , <i>rpl33</i> , <i>rpl36</i>
	RNA polymerase subunits/transcription RNA	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> *, <i>rpoC2</i>
	polymerase-associated factor	<i>pafl</i> *
RNA genes	Ribosomal RNA	<i>rrn4.5</i> , <i>rrn5</i> , <i>rrn16</i> , <i>rrn23</i>
	Transfer RNA	<i>TrnA-UGC</i> *, <i>trnC-GCA</i> *, <i>trnD-GUC</i> , <i>trnE-UUC</i> , <i>trnF-GAA</i> , <i>trnG-MCAU</i> , <i>trnG-GCC</i> , <i>trnG-UCC</i> *, <i>trnH-GUG</i> , <i>trnI-CAU</i> , <i>trnI-GAU</i> *, <i>trnK-UUU</i> *, <i>trnL-CAA</i> , <i>trnL-UAA</i> *, <i>trnL-UAG</i> , <i>trnM-CAU</i> , <i>trnN-GUU</i> , <i>trnP-UGG</i> , <i>trnQ-UUG</i> , <i>trnR-ACG</i> , <i>trnR-UCU</i> , <i>trnS-CGA</i> , <i>trnS-GCU</i> , <i>trnS-UGA</i> , <i>trnT-GGU</i> , <i>trnT-UGU</i> , <i>trnV-GAC</i> , <i>trnV-UAC</i> *, <i>trnW-CCA</i> , <i>trnY-GUA</i>
Other genes	Maturase	<i>matK</i>
	Protease	<i>clpP</i> *
	Acetyl-CoA carboxylase	<i>accD</i>
	Envelope membrane protein	<i>cemA</i>
	Cytochrome c synthesis	<i>ccsA</i>
Unknown function		<i>ycf2</i> , <i>ycf3</i> *, <i>ycf15</i>
Pseudogenes		<i>InfA</i> , <i>rps16</i> *, <i>ycf1</i>

The genes in bold are duplicated, the asterisk indicates genes with introns.



strong bias toward A/T homopolymers (A = 39, T = 44) (Figure 2A). Dinucleotide and trinucleotide repeats were equally represented (6 each, 6.2%), whereas only a single tetranucleotide repeats was detected (1.04%) (Figure 2B). The length distribution of mononucleotide SSRs exhibited a characteristic exponential decay, with a peak frequency of 17.4 at 10 bp, decreasing to 10.0 at 11 bp and rapidly declining rapidly to near zero beyond 15 bp (Figure 2C). This pattern confirms the evolutionary preference for shorter, more stable repeat motifs within chloroplast genomes.

Single-copy regions harbored the majority of SSRs, with the LSC region containing 73 SSRs at the highest density (79.81 per 100 kb), followed by the SSC region with 13 SSRs (71.10 SSRs per 100 kb). In contrast, the IR regions show remarkably low SSR content, with only five SSRs in each IR (18.75 SSRs per 100 kb). Across all genomic regions, mononucleotide repeats (p1) dominate all regions with 83 occurrences, followed by dinucleotide (p2) and trinucleotide (p3) repeats with 6 occurrences each. The single tetranucleotide repeat (p4) was exclusively located within the LSC region (Figure 3; Table 3).

3.2.2 Comparative analysis of SSRs across ten *Euphorbia* species: SSR compositional patterns and morphological differentiation

To assess the evolutionary and compositional significance of SSR patterns in *E. resinifera*, we examined chloroplast SSR characteristics across ten *Euphorbia* species to evaluate genus-wide trends and species-specific variations. *E. resinifera* occupies a central position on the total SSR density regression line ($r = -0.99$; Figure 4), showing an intermediate SSR abundance consistent with other cactiform species in the sect. *Euphorbia*.

Comparative analysis of adenine-type (A) and thymine-type (T) mononucleotide SSR distribution across complete chloroplast genomes (Figure 5) revealed distinct, species-specific compositional patterns within the genus. The pachycaul species *E. drupifera* uniquely exhibits perfect A=T equilibrium (50% each), while the annual herb *E. prostrata* shows near-equilibrium with a slight T enrichment 50.7%. In contrast, most congeners display a pronounced T-enrichment, particularly *E. neoarborescens*, which exhibits the highest SSR

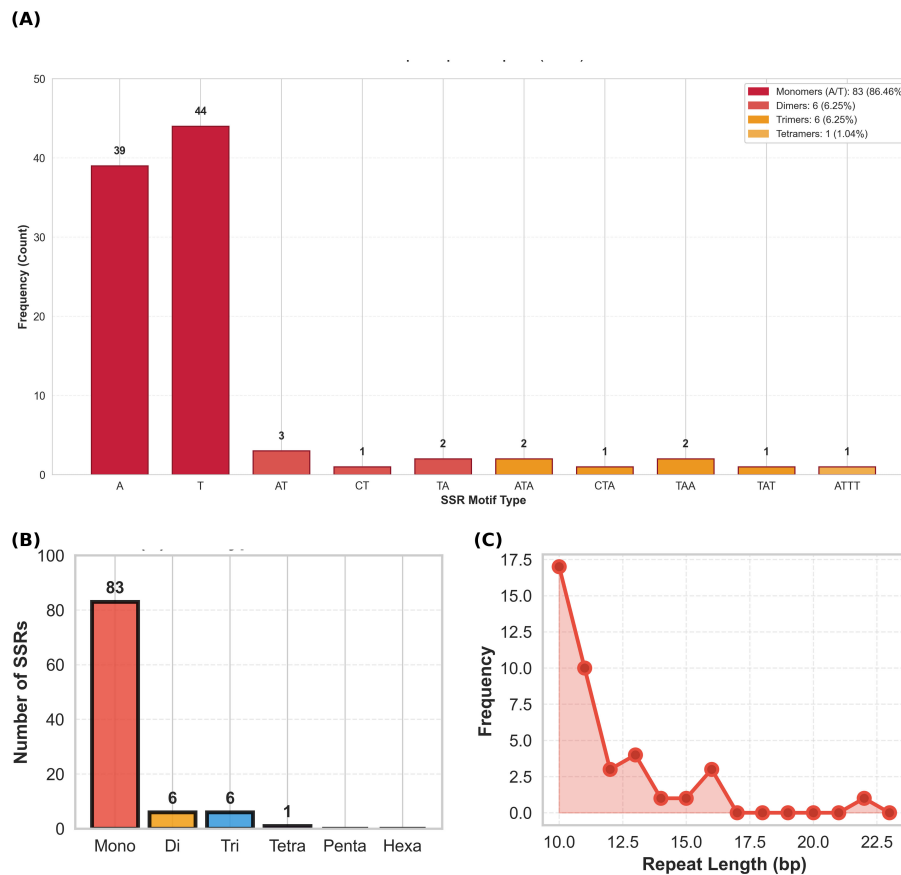


FIGURE 2

Simple sequence repeat analysis in *Euphorbia resinifera* chloroplast genome: (A) SSR type distribution, (B) marker classification, (C) mononucleotide length distribution.

abundance with strong T enrichment (64.2% T), and *E. lathyris* (64.3% T). *E. resinifera* shows a relatively balanced AT composition (47% A, 53% T), consistent with its intermediate SSR profile and providing a distinctive compositional signature within cactiform *Euphorbia*.

Analysis of the repetition patterns of simple sequence repeat motifs revealed multiple distinct motif classes according to their size, including monomers (A, T), dinucleotides (AT, CT, TA), trinucleotides (ATA, CTA, TAA, TAT, AAT, TTA, ATT, CAA),

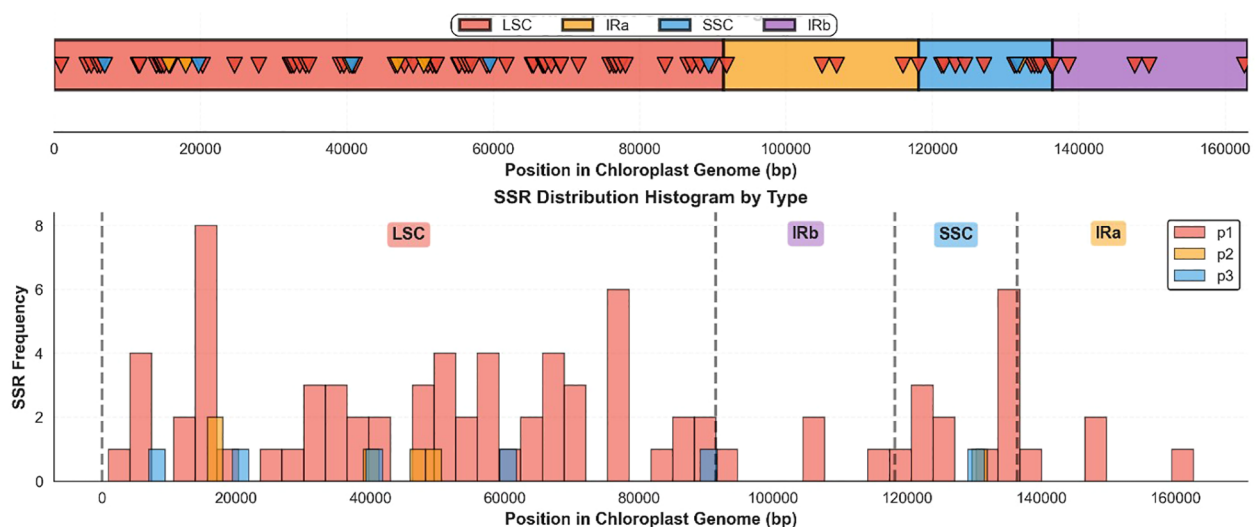


FIGURE 3

Distribution and density of simple sequence repeats (Mononucleotide (p1), Dinucleotide (p2), and Trinucleotide (p3), across the *Euphorbia resinifera* chloroplast genome regions.

TABLE 3 Quantitative analysis of SSR distribution across the quadripartite chloroplast genome structure of *Euphorbia resinifera*.

Region	Size (bp)	Number of SSRs	%	p1	p2	p3	p4	Density/100kb
LSC	91,462	73	76.04	62	5	5	1	79.81
SSC	18,285	13	13.54	11	1	1	0	71.10
IRa	26,659	5	5.208	5	0	0	0	18.75
IRb	26,659	5	5.208	5	0	0	0	18.75

tetranucleotides (ATTT), as well as pentanucleotide and hexanucleotide motifs (Supplementary Figure 1). Sørensen-Dice analysis identified TTA as the most phylogenetically informative motif (index = 0.5, present in 5/10 species), followed by CTA, TAA, AAT, and TATAT (index = 0.3) (Table 4). The predominance of A/T-rich motifs reflects the characteristic composition of chloroplast genomes.

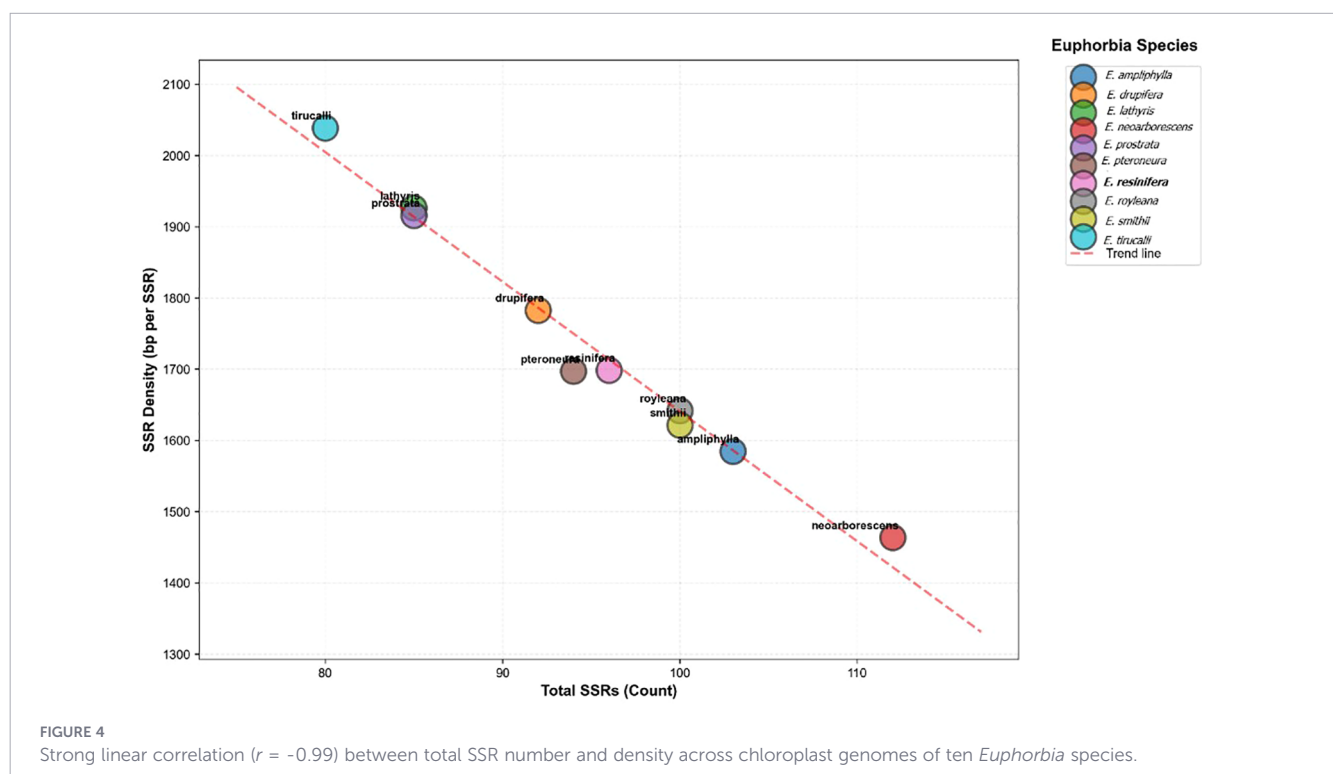
The Sørensen dissimilarity matrix reveals three SSR-based compositional clusters with low within cluster dissimilarities (0.14–0.45) and higher between-cluster distances (0.33–0.47). Cluster I groups all four species in sect. *Euphorbia* (*E. resinifera*, *E. ampliphylla*, *E. drupifera*, and *E. royleana*), with *E. resinifera* and *E. ampliphylla* showing highest similarity (0.14) (Figure 6; Supplementary Table 1). However, Clusters II and III mix species across subgenera: Cluster II associates *E. neoarborescens* and *E. tirucalli* (subgenus *Euphorbia*) with *E. smithii* (subgenus *Athymalus*), while Cluster III groups *E. prostrata* (subgenus *Chamaesyce*), *E. lathyris* (subgenus *Esula*), and *E. pteroneura* (subgenus *Euphorbia*). Thus, cpSSR compositional profiles provide taxonomically informative characters exclusively for resolving sectional monophyly in sect. *Euphorbia*, while lacking phylogenetic signal for higher-order taxonomic relationships due to extensive character state convergence and evolutionary homoplasy.

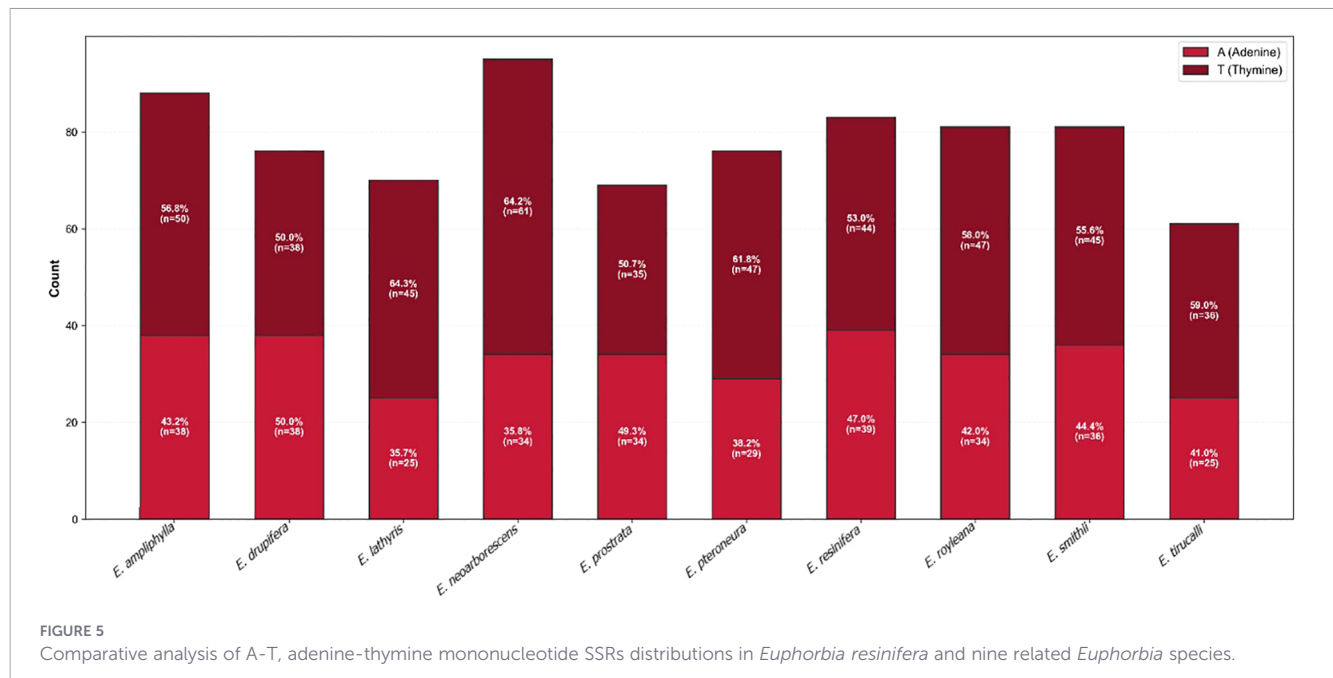
Analysis of the T/A ratio gradient shows gradual enrichment in T bases, progressing from the balanced pachycaul species *E. drupifera* (T/A = 1.00) through increasingly T-enriched cactiform species: *E. resinifera* (1.13), *E. ampliphylla* (1.32) and *E. royleana* (1.38) (Figure 7).

3.3 Phylogenetic relationships and molecular resolution in sect. *Euphorbia*

3.3.1 Maximum likelihood phylogenetic reconstruction

The maximum likelihood (ML) phylogenetic tree shows that *E. resinifera* clusters with other *Euphorbia* species, forming a monophyletic group with maximal bootstrap support (BP = 100). *E. resinifera* is particularly closely related to *E. ampliphylla*, *E. drupifera*, and *E. royleana*, all of which share a common ancestor within sect. *Euphorbia*. However, *E. drupifera* is a non-cactiform species, unlike *E. resinifera*, *E. ampliphylla* and *E. royleana*, which are xerophytic cactiform species. It is also noteworthy that *E. resinifera*, *E. ampliphylla* and *E. drupifera* are distributed in Africa, whereas *E. royleana* is native to Asia (Figure 8). Overall, these results further suggest a close evolutionary relationship





among the 22 *Euphorbia* species, which cluster together on a strongly supported branch (BP = 100%). Moreover, the maximal bootstrap support for the *E. resinifera* and *E. ampliphylla* clade indicates strong confidence in their inferred relationship.

3.3.2 IR/SC boundary analysis in *E. resinifera* and related species

Comparative analysis of IR/SC boundary regions across 20 *Euphorbia* species revealed substantial interspecific variation in the positioning, orientation, and extension of key genes (*rps19*, *rpl22*, *rpl2*, *trnH*, *ycf1*, and *ndhF* ($P_i = 0.0091$)) at the four junctions

(LSC/IRb, IRb/SSC, SSC/IRa, IRa/LSC). Within section *Euphorbia* (*E. resinifera*, *E. ampliphylla*, *E. drupifera*, and *E. royleana*), differences were minimal with junction shifts of only 74 bp in the IR, 155 bp in the SSC, and 1,120 bp in the LSC. Among these regions, the *rpl22/rps19* boundary at the LSC/IRb junction exhibited the highest variability. *E. resinifera*, *E. ampliphylla*, and *E. royleana* shared identical *rpl22* positioning (357 bp in the LSC and 66 bp in the IRb), whereas *E. drupifera* showed a slight expansion (360 bp in the LSC, 66 bp in the IRb). Despite these minor differences in the LSC portion, *rpl22* consistently extended 66 bp into the IRb region across all four species within section *Euphorbia*, indicating a conserved structural feature characteristic of this taxonomic group (Figure 9).

TABLE 4 Top phylogenetically informative SSR motifs.

Motif	Present in species	Absent in species	Sørensen-dice index
TTA	5	5	0.5
CTA	3	7	0.3
TAA	7	3	0.3
AAT	7	3	0.3
TATAT	3	7	0.3
ATTT	2	8	0.2
ATATAC	2	8	0.2
AAATAG	2	8	0.2
ATA	9	1	0.1
TAT	9	1	0.1
AAAGA	1	9	0.1
TTATA	1	9	0.1
TGTATA	1	9	0.1
TATAG	1	9	0.1
TAAAA	1	9	0.1

3.3.3 Nucleotide polymorphism distribution identifies informative markers

Comparative nucleotide diversity analysis of 129 chloroplast genes across 22 *Euphorbia* species representing four subgenera reveals uneven distribution of hypervariable regions (HVRs) (Figure 10; Table 5). The top 13 most variable genes are exclusively LSC-encoded with P_i values ranging from 0.081 (*trnH*-GUG) to 0.046 (*matK*), while 32 highly conserved genes exhibit minimal polymorphism ($P_i \leq 0.012$), including *ndhF* ($P_i = 0.009$), *trnL*-UAG ($P_i = 0.008$), and *trnL*-CAA ($P_i = 0.007$), with three tRNA loci showing complete conservation ($P_i = 0.000$). Notably, *rpl22* ($P_i = 0.057$, rank 5) exhibits 23% higher nucleotide diversity than traditional markers *matK* ($P_i = 0.046$), *rbcl* ($P_i = 0.034$), and *ndhF* ($P_i = 0.0091$) (Supplementary Table 3).

Regional distribution analysis reveals marked variation across chloroplast genome compartments. The LSC region shows the highest evolutionary rate with 59 of 81 genes (72.8%) classified as variable to highly variable, while Inverted Repeats exhibit strong conservation with only 3 of 18 genes (16.7%) showing variability in each IR copy, and the SSC region contains 12 entirely conserved

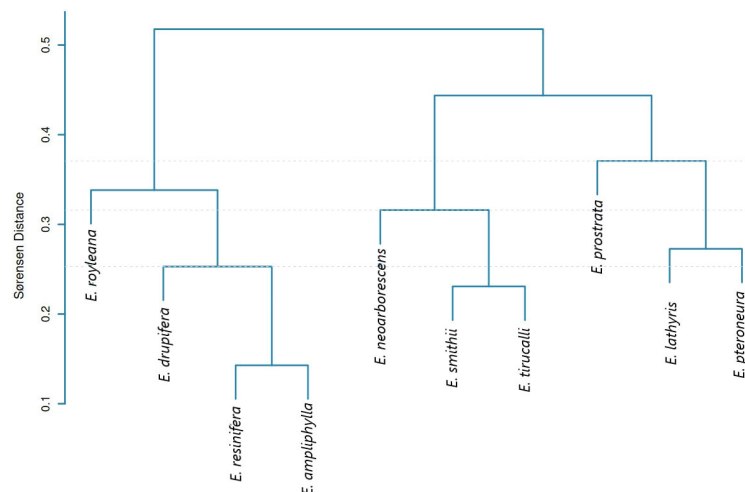


FIGURE 6

Sørensen distance-based dendrogram showing SSR compositional clustering with taxonomic signal limited to sectional monophyly in sect. *Euphorbia*, while subgeneric relationships are confounded by homoplasy.

genes (0.0% variability). Overall, 65 of 129 chloroplast genes (50.4%) exhibited variability ($P_i \geq 0.024$).

3.3.4 *rpl22*-based phylogenetic resolution of morphological groups

Analysis of the *rpl22* gene reveals two distinct groups: *E. ampliphylla* and *E. resinifera* exhibit 100% sequence identity and share 99.76% similarity with *E. royleana*, while *E. drupifera* displays 2.6% variation (97.41–97.87% similarity) characterized by 10 SNPs and retention of a unique AGC trinucleotide at positions 395–397 that is deleted in the other three species. Nine SNPs (positions 223, 385, 390, 391, 392, 394, 396, 397, 409) are identical across the three cactiform species (*E. resinifera*, *E. ampliphylla*, *E. royleana*), indicating a single divergence event from the pachycaul *E. drupifera*. Position 405 represents a biogeographical marker distinguishing Asian *E. royleana* (G) from African taxa (T), confirming biogeographical structure within the otherwise highly conserved gene (96.95% overall identity) (Figure 11; Supplementary

Figure 3). Six mutations align with T/A enrichment through adenine loss (A→G at positions 223, 385, 390; A→T at positions 392, 396) and thymine gain (G→T at position 394), with A→T transversions showing the strongest T/A enrichment effect. Transitions and transversions contribute equally (5 mutations each) (Table 6).

Phylogenetic analysis of the *rpl22* gene (306–537 bp) from 22 species confirms the monophyly of *Euphorbia* with maximum bootstrap support (BS = 100%), covering the subgenus *Euphorbia* (green, BS = 97%), the subgenus *Chamaesyce* (blue, BS = 100%) and the subgenus *Esula* (red, BS = 70%). Within the subgenus *Euphorbia*, the section *Euphorbia* includes the three cactiform species *E. ampliphylla*, *E. resinifera*, and *E. royleana* (BS = 85%), with *E. ampliphylla* and *E. resinifera* being identical over 423 bp (0 SNPs) and *E. royleana* differing by a single SNP. The *Monadenium* section is well resolved (BS = 96%), with *E. neospinescens* and *E. neoborens* showing 100% sequence identity over 444 bp, while the herbaceous *E. neogossweileri* appears as their sister taxon (BS = 70%). The subgenus *Chamaesyce* is the best-supported clade

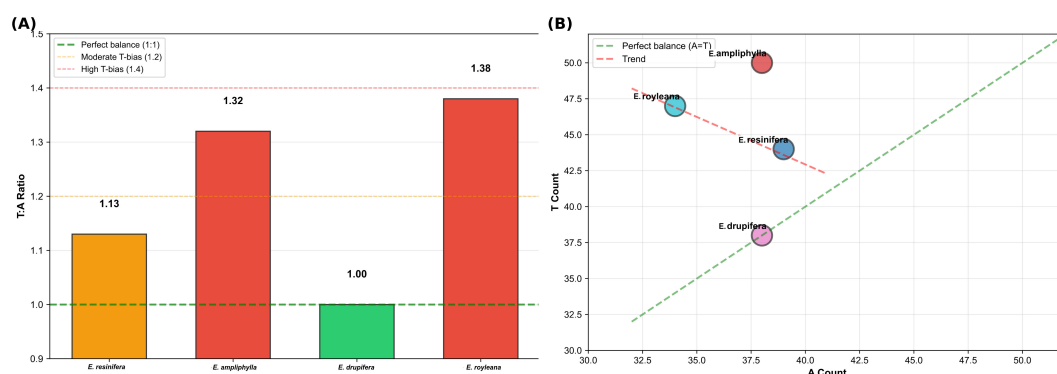


FIGURE 7

Comparative differentiation of chloroplast SSR monomers among four *Euphorbia* species (sect. *Euphorbia*). (A) T/A ratio gradient showing progressive thymine enrichment from pachycaul to cactiform species. (B) A-T scatter plot highlighting compositional divergence and revealing molecular basis for taxonomic discrimination between cactiform and pachycaul lineage.

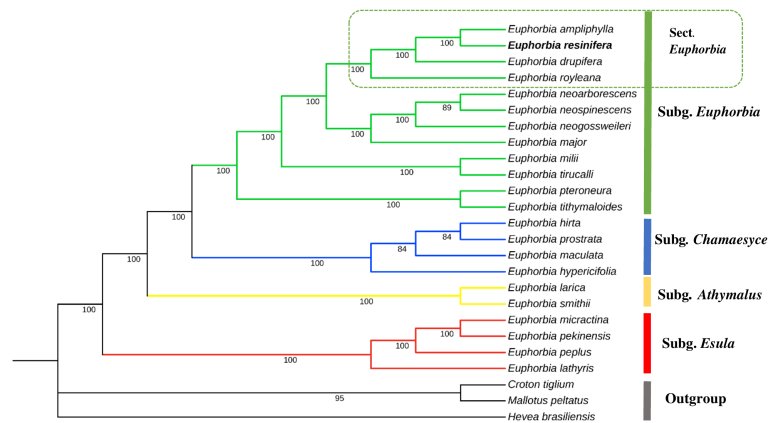


FIGURE 8

Maximum likelihood (ML) tree based on the chloroplast sequences of 25 species, including 22 *Euphorbia* species and three Euphorbiaceae outgroups (*Hevea brasiliensis*, *Croton tiglium*, *Mallotus peltatus*). The sequences incorporated in the tree includes *E. resinifera* (this study) and 24 publicly available chloroplast genomes.

in the tree (BS = 100%). The most notable result concerns *E. drupifera*, whose pachycaul morphology had previously obscured its taxonomic position; the *rpl22* gene excludes it from the sect. *Euphorbia* and resolves its position with *E. milii* and *E. tirucalli* (BS = 76%). The subgenus *Esula* has limited internal resolution, with only *E. micractina* and *E. pekinensis* forming a well-supported clade (BS = 94%), these two species being 100% identical over 306 bp, while the other species appear as isolated branches (Figure 12).

4 Discussion

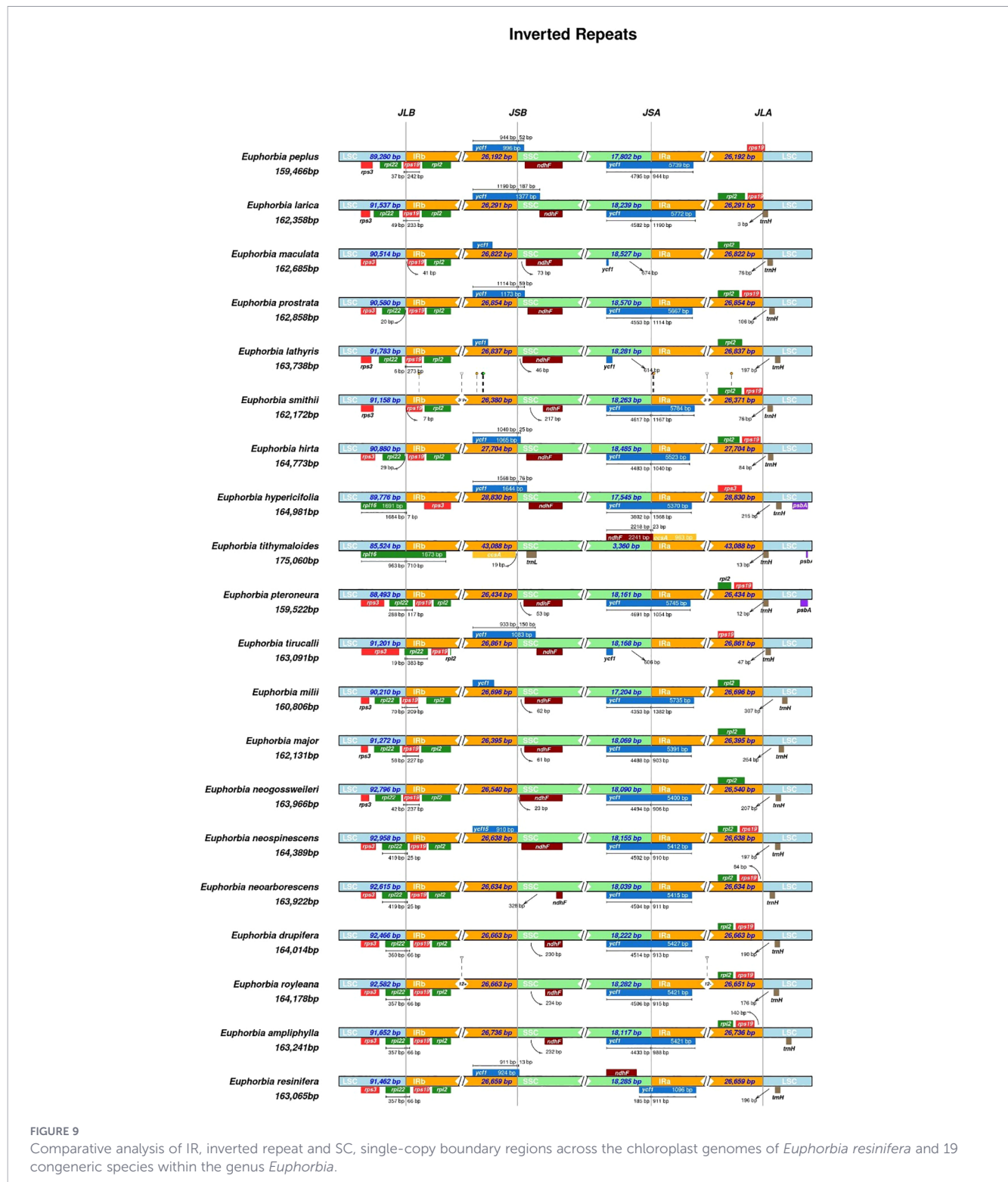
The *Euphorbia resinifera* chloroplast genome (163,065 bp, 35.11% GC, 132 genes) exhibits typical quadripartite structure within the ranges reported for *Euphorbia* species by Lee et al. (2024) and Wei et al. (2021), demonstrating remarkable structural conservation across the genus. The low GC content (35.11%) matches other subgenus *Euphorbia* species (*E. drupifera*, *E. neogossweileri*, *E. tithymaloides*) reported by Wei et al. (2021), supporting the phylogenetic coherence of this taxonomic section and providing valuable baseline data for future comparative genomic studies.

Simple sequence repeat analysis revealed strong A/T bias, with mononucleotide repeats comprising 86.5% of all detected SSRs and A/T homopolymers predominating (A = 39, T = 44), consistent with widespread patterns in land plant plastomes where A/T-type SSRs frequently exceed 90% (Kan et al., 2024), reflecting characteristic chloroplast genome composition (Kuang et al., 2011; Asaf et al., 2018; Xu et al., 2025) and attributable to polyadenylation processes and the relative ease of A/T strand separation during replication increasing slippage probability (Dodsworth et al., 2015). The AT-rich regions are often associated with easier unwinding of DNA during transcription and potentially more efficient and accurate translation processes (Wang et al., 1984). The prevalence of these A/T-rich repetitive

sequences is thought to contribute to genomic plasticity and stability, with their specific motifs and distribution patterns likely reflecting evolutionary pressures that shaped plant genomes for adaptation and divergence (Wu et al., 2018; Dong et al., 2025). SSR density is approximately four times higher in single-copy regions than in inverted repeats, consistent with substitution rates 3.7 times slower in IRs (Zhu et al., 2016), while IR boundary contraction and expansion events drive interspecific genome size variation (Shen et al., 2017; Corvalán et al., 2023). While cpSSR compositional profiles successfully recover sect. *Euphorbia* as a distinct cluster, their failure to resolve established subgeneric boundaries in Clusters II and III confirms the well-documented limitations of microsatellite markers for deep phylogenetic inference. Although microsatellite markers are prone to homoplasy (Kuang et al., 2011; Dodsworth et al., 2015; Vieira et al., 2016) this limitation has been addressed by incorporating whole-chloroplast sequencing to directly examine underlying sequence variation (Wheeler et al., 2014).

Complete chloroplast genome analyses confirm the monophyly of the four *Euphorbia* subgenera (*Esula*, *Chamaesyce*, *Athymalus*, and *Euphorbia*) and place *E. resinifera* within section *Euphorbia*, together with *E. ampliphylla*, *E. drupifera*, and *E. royleana*, in agreement with previous taxonomic classifications (Horn et al., 2012; Dorsey et al., 2013; Wei et al., 2021). Within sect. *Euphorbia*, the two African cactiform species (*E. ampliphylla* and *E. resinifera*) form a sister group, and together with the pachycaul *E. drupifera* constitute a well-resolved African clade, with the Asian *E. royleana* occupying a basal position relative to this clade, demonstrating that biogeographic origin strongly influences plastome evolution. However, the grouping of the pachycaul *E. drupifera* with African cactiform species in the complete plastome analysis, despite its fundamentally distinct morphology, illustrates how shared biogeographic origin can override true phylogenetic signal, consistent with the well-documented influence of geography on chloroplast genome evolution (He et al., 2024).

A key architectural distinction characterizes *Euphorbia* species within the section *Euphorbia*: atypical positioning of the *rpl22* gene



spanning the LSC/IRb junction (357–360 bp in LSC and 66 bp in IRb), in contrast to its conventional localization entirely within the LSC in most angiosperm plastomes (Meng et al., 2018; Ahmed and Rahman, 2024). Similar boundary variations have been documented in *Paphiopedilum* orchids, where *rpl22* extends into both regions in 6 of 14 species (Guo et al., 2021). This configuration represents a stable synapomorphy among cactiform taxa (*E. resinifera*, *E. ampliphylla*, and *E. royleana*), whereas the pachycaul species

E. drupifera exhibits slight boundary modification (360 bp in LSC/ 66 bp in IRb), indicating a subtle structural divergence. This unique positioning subjects *rpl22* to distinct evolutionary dynamics, potentially influenced by the homogenizing effects of gene conversion in the IR region while maintaining sufficient variation in the LSC portion, consequently offering effective infrageneric resolution by balancing variability and conservation (Nie et al., 2025).

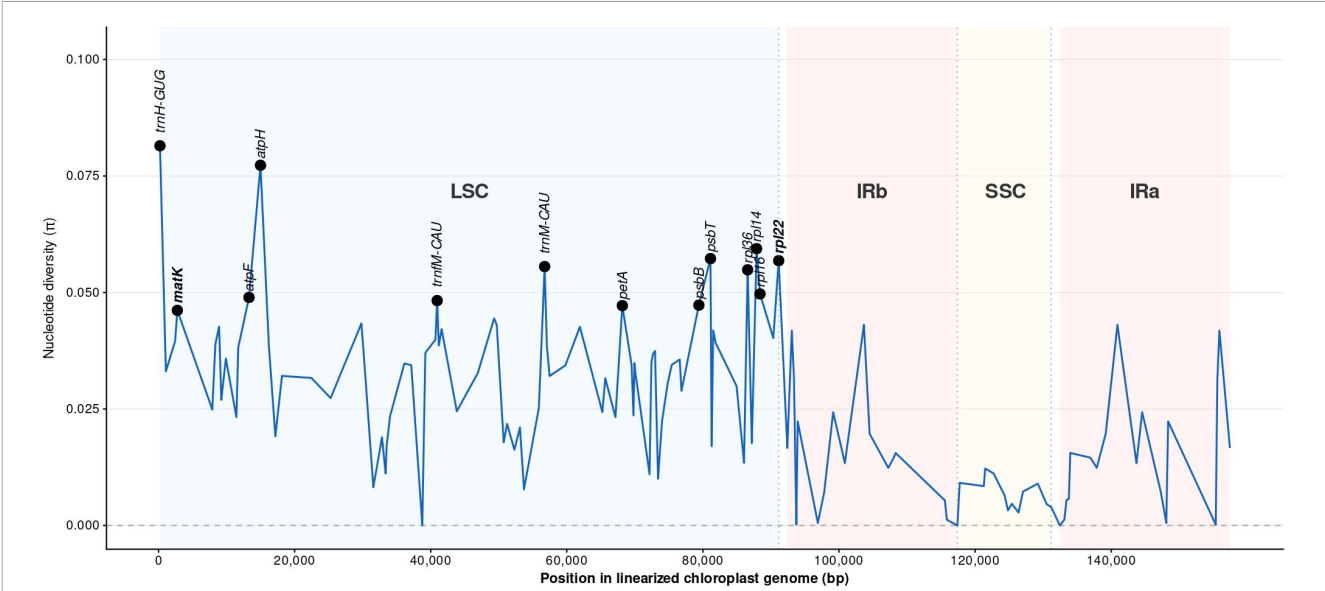


FIGURE 10 Nucleotide diversity (Pi) across chloroplast genomes of 22 *Euphorbia* species by sliding window analysis, highlighting the top 13 most variable genes.

Regional polymorphism analysis highlights the limited suitability of conventional plastid markers for resolving fine-scale phylogenetic relationships in *Euphorbia*. Compartment-specific evolutionary patterns reveal pronounced LSC dominance in sequence variation, while IR regions maintain strong conservation through gene conversion mechanisms, with highly variable regions (HVRs) predominantly concentrated in non-coding sequences, similar to patterns observed across angiosperms (Shi et al., 2023; Xu et al.,

2025). This evolutionary heterogeneity reflects fundamental differences in selective constraints, where IR-SC junctions remain highly conserved with stable boundary markers (*rps19*, *ycf1*, *ndhF*, *trnN*, *trnH*), contrasting sharply with LSC-driven variability.

Despite this compartment-specific variation, widely used phylogenetic markers demonstrate variable and often insufficient discriminatory power for species-level resolution. While some markers like *matK* exhibit high variability, others including *rbcl*,

TABLE 5 Distribution of variable genes across chloroplast genome regions in 22 *Euphorbia* species.

Category (pi range)	LSC	SSC	IR	Total	% of total
Highly variable (0.038 - 0.081)	29	0	4	33	25.6%
Variable (0.024 - 0.037)	30	0	2	32	24.8%
Conserved (0.012 - 0.024)	16	0	16	32	24.8%
Highly conserved (0.000 - 0.012)	6	12	14	32	24.8%
Total	81	12	36	129	100%

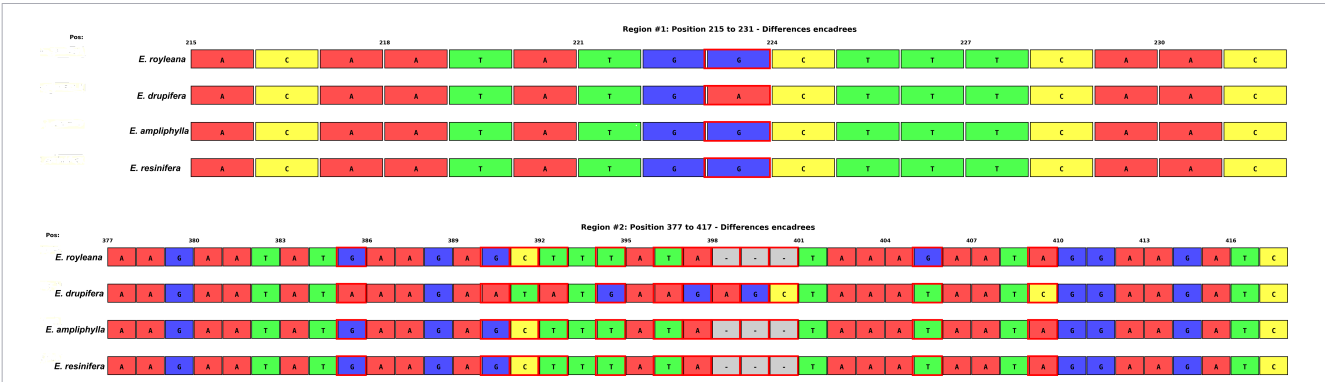


FIGURE 11 Multiple sequence alignment of the 426-bp *rpl22* gene showing 10 variable sites and a 3-bp indel distinguishing lineages within section *Euphorbia*.

TABLE 6 Single nucleotide polymorphisms (SNPs) in the *rpl22* gene among four section *Euphorbia* species: comparative analysis of three cactiform and one pachycaul morphotype.

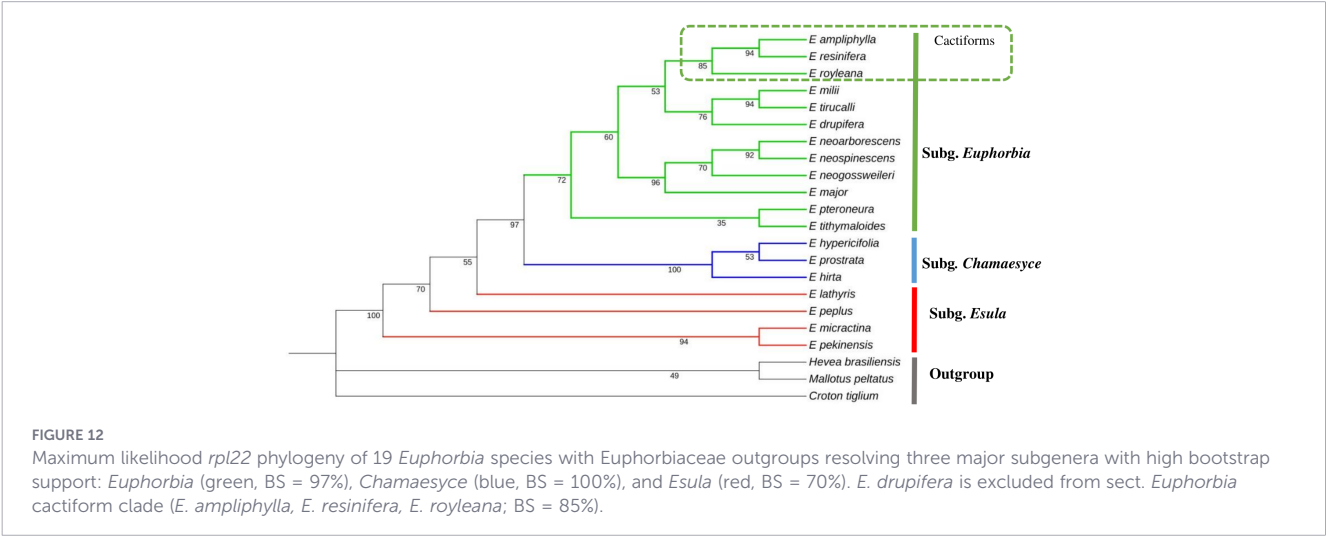
Position	Type	Nucleotide change	<i>E. drupifera</i>	<i>E. resinifera</i>	<i>E. ampliphylla</i>	<i>E. royleana</i>	T/A effect
223	Transition	A→G	A	G	G	G	Loses A
385	Transition	A→G	A	G	G	G	Loses A
390	Transition	A→G	A	G	G	G	Loses A
391	Transition	T→C	T	C	C	C	Loses T
397	Transition	G→A	G	A	A	A	Gains A
392	Transversion	A→T	A	T	T	T	Loses A, Gains T
394	Transversion	G→T	G	T	T	T	Gains T
396	Transversion	A→T	A	T	T	T	Loses A, Gains T
409	Transversion	C→A	C	A	A	A	Gains A
405	Transversion	T→G	T	T	T	G	Loses T

psbA, *ndhF*, *trnL*-UAG, and *trnL*-CAA show limited phylogenetic informativeness, with multiple studies demonstrating inadequate resolution for discriminating closely related species in taxonomically complex genera (Muellner et al., 2003; Shi et al., 2023; Nie et al., 2025). This systematic limitation necessitates identification of alternative loci with enhanced phylogenetic utility. In contrast, *rpl22* exhibits exceptional species-level discriminatory utility, aligning with comparative genomics studies identifying *rpl22* and its flanking sequences as evolutionarily dynamic mutation hotspots suitable for DNA barcoding applications (Wang et al., 2021, 2024). The documented positive selection in *Euphorbia altotibetica* (Qu et al., 2025) and *Cistanche* species (Miao et al., 2022) indicates rapid adaptive evolution in this ribosomal protein L22-encoding gene, suggesting functional significance for maintaining protein synthesis and stress adaptation in arid environments. This combination of evolutionary dynamism and potential functional adaptation establishes *rpl22* as a promising phylogenetic marker for systematic studies in xerophytic lineages where traditional plastid markers lack sufficient resolution.

Phylogenetic analysis based on Maximum Likelihood reconstruction demonstrates that the *rpl22* gene better reflects

morphological relationships than complete chloroplast genome phylogeny. While whole-plastome analyses cluster species by geographic origin, *rpl22* correctly excludes *E. drupifera* from sect. *Euphorbia* based on molecular grounds consistent with its distinct pachycaul morphology. As noted by Dorsey et al. (2013), although most species of sect. *Euphorbia* possess spiny succulent stems, leafy trees such as *E. drupifera* represent morphological exceptions, with persistent leaves interpreted as ancestral state reversal (Horn et al., 2012).

The 2.6% molecular distinction of *E. drupifera* (*rpl22*) and its T/A ratio (cpSSRs) of 1.00, contrasting with values of 1.13–1.38 in cactiform species, may correlate with fundamental photosynthetic strategy differences between stem-dominated and leaf-dominated growth forms (Horn et al., 2012). Remarkable sequence conservation among morphologically similar species (0–1 SNPs among cactiform species of 423 bp; 100% identity between *E. neospinescens* and *E. neoarborescens* of 444 bp and *E. micractina* and *E. pekinensis* of 306 bp) demonstrates that *rpl22* reflects deep evolutionary relationships rather than recent biogeographic history. These findings highlight the value of targeted single-gene approaches for resolving systematic relationships in hyperdiverse genera



where geographic clustering can confound whole-plastome phylogenetic inference.

5 Conclusion

This study demonstrates that integrating multiple independent molecular approaches provides robust taxonomic resolution in *Euphorbia*, overcoming individual methodological limitations including microsatellite homoplasy. Using a novel non-destructive spine-DNA genome-skimming protocol, we generated the first complete chloroplast genome of the endemic arid-adapted cactiform *E. resinifera* from North-West African hyperarid zones. Our findings establish *rpl22* as a superior phylogenetic marker that differentiates cactiform xerophytes from pachycaul species through lineage-specific substitution patterns and diagnostic SNPs correlating with photosynthetic architectural transitions, distinctions undetectable using conventional plastid markers (*rbcL*, *matK*, *ndhF*). The *rpl22* gene resolves taxonomic boundaries congruent with whole-genome classifications while achieving maximum bootstrap support for *Euphorbia* relationships. Concatenation with additional genes is needed to enhance phylogenetic robustness and remove evolutionary ambiguities in taxonomic relationships.

Data availability statement

Data are deposited in: Zenodo: <https://doi.org/10.5281/zenodo.18750510> (gene lists, phylogenetic files, CDS verification, pseudogenes, tRNA/rRNA data). NCBI GenBank: BioProject PRJNA1371610, BioSample SAMN53505053, accession number PZ168867 (BankIt3063561).

Author contributions

AT: Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. KR: Writing – review & editing, Data curation. ML: Data curation, Writing – review & editing, Supervision. FG: Formal analysis, Visualization, Writing – original draft. IL: Data curation, Investigation, Methodology, Writing – review & editing. NL: Investigation, Writing – original draft. GD: Validation, Writing – original draft. RA: Validation, Writing – original draft. KS: Investigation, Writing – original draft. AE: Investigation, Resources, Writing – original draft. ME: Investigation, Visualization, Writing – original draft. HK: Validation, Writing – original draft. MJ: Conceptualization, Visualization, Writing – review & editing. MF: Conceptualization, Writing – review & editing, Funding acquisition, Project administration, Supervision. YA: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. Funds were provided by Texas Tech University and the Fulbright program, which provided the necessary funding to cover research expenses for MF.

Acknowledgments

MF gratefully acknowledges financial support for this project by the Fulbright U.S. Scholar Program, which is sponsored by the U.S. Department of State and the Moroccan-American Commission for Educational and Cultural Exchange (MACECE). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Fulbright Program, the Government of the United States, or MACECE.

Conflict of interest

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

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2026.1785579/full#supplementary-material>

References

- Ahmed, S. S., and Rahman, M. O. (2024). Deciphering the complete chloroplast genome sequence of *Meconopsis torquata* Prain: Insights into genome structure, comparative analysis and phylogenetic relationship. *Heliyon* 10, e36204.
- Al-Qthanin, R. N., and Alharbi, S. A. (2023). Regional conservation assessment of the threatened species: A case study of twelve plant species in the Farasan Archipelago. *Conservation* 3, 127–152. doi: 10.3390/conservation3010011
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> (Accessed July 16, 2025).
- Asaf, S., Khan, A. L., Khan, M. A., Shahzad, R., Lubna, Kang, S. M., et al. (2018). Complete chloroplast genome sequence and comparative analysis of loblolly pine (*Pinus taeda* L.) with related species. *PLoS One* 13, e0192966. doi: 10.1371/journal.pone.0192966
- Beier, S., Thiel, T., Münch, T., Scholz, U., and Mascher, M. (2017). MISA-web: A web server for microsatellite prediction. *Bioinformatics* 33, 2583–2585. doi: 10.1093/bioinformatics/btx198
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Bruyns, P. V., Mapaya, R. J., and Hedderson, T. J. (2006). A new subgeneric classification for *Euphorbia* (Euphorbiaceae) in southern Africa based on ITS and psbA-trnH sequence data. *Taxon* 55, 397–420. doi: 10.2307/25065587
- Chaudhary, P., Singh, D., Swapnil, P., Meena, M., and Janmeda, P. (2023). *Euphorbia nerifolia* (Indian Spurge Tree): A plant of multiple biological and pharmacological activities. *Sustainability* 15, 1225. doi: 10.3390/su15021225
- Corvalán, L. C. J., Sobreiro, M. B., Carvalho, L. R., Dias, R. O., Braga-Ferreira, R. S., Targueta, C. P., et al. (2023). Chloroplast genome assembly of *Serjania erecta* Raldk: comparative analysis reveals gene number variation and selection in protein-coding plastid genes of Sapindaceae. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1258794
- Dierckxens, N., Mardulyn, P., and Smits, G. (2020). Unraveling heteroplasmy patterns with NOVOPlasty. *NAR Genomics Bioinform.* 2, lqz011. doi: 10.1093/nargab/lqz011
- Díez Menéndez, C., Pocai, P., Williams, B., Myllys, L., and Amirouf, A. (2023). IRplus: an augmented tool to detect inverted repeats in plastid genomes. *Genome Biol. Evol.* 15, evad177. doi: 10.1093/gbe/evad177
- Dodsworth, S., Chase, M. W., Kelly, L. J., Leitch, I. J., Macas, J., Novak, P., et al. (2015). Genomic repeat abundances contain phylogenetic signal. *Syst. Biol.* 64, 112–126. doi: 10.1093/sysbio/syu080
- Dong, Y., Cao, Q., Yu, K., Wang, Z., Chen, S., Chen, F., et al. (2025). Chloroplast phylogenomics reveals the maternal ancestry of cultivated *chrysanthemums*. *Genomics Commun.* 2, e019. doi: 10.48130/gcomm-0025-0019
- Dorsey, B. L., Haevermans, T., Aubriot, X., Morawetz, J. J., Riina, R., Steinmann, V. W., et al. (2013). Phylogenetics, morphological evolution, and classification of *Euphorbia* subgenus *Euphorbia*. *Taxon* 62, 291–315. doi: 10.12705/622.1
- Ettaqy, A., Taha, A., ElGhiouane, A., ElKhou, A., Boulli, A., and Abbas, Y. (2020). “New data on the ecological distribution of *Euphorbia resinifera* O. Berg in the Beni Mellal-Khenifra region,” in *E3S Web Conf*, 183, 01001. doi: 10.1051/e3sconf/202018301001
- Fehlberg, S. D., Allen, J. M., and Church, K. (2013). A novel method of genomic DNA extraction for Cactaceae. *Appl. Plant Sci.* 1, 10–13. doi: 10.3732/apps.1200013
- Fennane, M. (2018). *Éléments pour un Livre rouge de la flore vasculaire du Maroc. Fasc. 10. Monocotyledonae. Tela-Botanica* 10, 1–92. Available online at: https://www.tela-botanica.org/wp-content/uploads/2018/01/Tela-Bot_LivreR-FVM_Fasc-5-janv.-2018.pdf (Accessed October 8, 2025).
- Greiner, S., Lehwark, P., and Bock, R. (2019). OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* 47, W59–W64. doi: 10.1093/nar/gkz238
- Guo, B., Zeng, S., Yin, Y., Li, L., Ma, G., Wu, K., et al. (2021). Characterization of phytohormone and transcriptome profiles during protocorm-like bodies development of *Paphiopedilum*. *BMC Genomics* 22, 1–17. doi: 10.1186/s12864-021-08087-y
- He, M., Han, X., Qin, X., Bao, J., Li, H., Xie, Q., et al. (2024). Comparative chloroplast genome analyses provide new insights into phylogeny of *Taraxacum* and molecular markers for distinguishing rubber producing dandelions from their weedy relatives in China. *Ind. Crops Prod.* 207, 117712. doi: 10.1016/j.indcrop.2023.117712
- Horn, J. W., van Ee, B. W., Morawetz, J. J., Riina, R., Steinmann, V. W., Berry, P. E., et al. (2012). Phylogenetics and the evolution of major structural characters in the giant genus *Euphorbia* L. (Euphorbiaceae). *Mol. Phylogenet. Evol.* 63, 305–326. doi: 10.1016/j.ympev.2011.12.022
- Horn, J. W., Xi, Z., Riina, R., Peirson, J. A., Yang, Y., Dorsey, B. L., et al. (2014). Evolutionary bursts in *Euphorbia* (Euphorbiaceae) are linked with photosynthetic pathway. *Evolution* 68, 3485–3504. doi: 10.1111/evo.12534
- Jin, J. J., Yu, W. B., Yang, J. B., Song, Y., Depamphilis, C. W., Yi, T. S., et al. (2020). GetOrganelle: A fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biol.* 21, 241. doi: 10.1186/s13059-020-02154-5
- Kan, J., Nie, L., Wang, M., Tiwari, R., Tembrock, L. R., and Wang, J. (2024). The Mendelian pea pan-plastome: insights into genomic structure, evolutionary history, and genetic diversity of an essential food crop. *Genomics Commun.* 1, e004. doi: 10.48130/gcomm-0024-0004
- Katoh, K., Rozewicki, J., and Yamada, K. D. (2018). MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108
- Kuang, D. Y., Wu, H., Wang, Y. L., Gao, L. M., Zhang, S. Z., and Lu, L. (2011). Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): implication for DNA barcoding and population genetics. *Genome* 54, 663–673. doi: 10.1139/g11-026
- Lee, S. R., Oh, A., and Son, D. C. (2024). Characterization, comparison, and phylogenetic analyses of chloroplast genomes of *Euphorbia* species. *Sci. Rep.* 14, 16759. doi: 10.1038/s41598-024-66102-0
- Letunic, I., and Bork, P. (2024). Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* 52, W78–W82. doi: 10.1093/nar/gkac268
- Li, F., Xie, X., Huang, R., Tian, E., Li, C., and Chao, Z. (2021). Chloroplast genome sequencing based on genome skimming for identification of Eriobotryae Folium. *BMC Biotechnol.* 21, 69. doi: 10.1186/s12896-021-00728-0
- Masrahi, Y. S., Al-Turki, T. A., and Sayed, O. H. (2015). Photosynthetic adaptation and survival strategy of *Duvalia velutina* in an extremely arid environment. *Photosynthetica* 53, 555–561. doi: 10.1007/s11099-015-0143-0
- Meng, J., Li, X., Li, H., Yang, J., Wang, H., and He, J. (2018). Comparative analysis of the complete chloroplast genomes of four aconitum medicinal species. *Molecules* 23, 1015. doi: 10.3390/molecules23051015
- Miao, Y., Chen, H., Xu, W., Yang, Q., Liu, C., and Huang, L. (2022). Structural mutations of small single copy (SSC) region in the plastid genomes of five *Cistanche* species and inter-species identification. *BMC Plant Biol.* 22, 319. doi: 10.1186/s12870-022-03682-x
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., et al. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015
- Muellner, A. N., Samuel, R., Johnson, S. A., Cheek, M., Pennington, T. D., and Chase, M. W. (2003). Molecular phylogenetics of Meliaceae (Sapindales) based on nuclear and plastid DNA sequences. *Am. J. Bot.* 90, 471–480. doi: 10.3732/ajb.90.3.471
- Nie, Z., Ma, J., Wang, C., Tang, M., Jia, T., Liao, G., et al. (2025). Comparative analysis of chloroplast genomes on Meliaceae species: insights into the evolution and species identification. *Front. Plant Sci.* 16. doi: 10.3389/fpls.2025.1536313
- Qu, R., Mao, X., Zhou, Y., Liu, Y., Su, X., Hu, X., et al. (2025). The complete chloroplast genome and phylogeny of *Euphorbia altotibetica* (Euphorbiaceae), an endemic medicinal plant species from northwest China. *Pakistan J. Bot.* 57, 645–660. doi: 10.30848/PJB2025-2(33)
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904. doi: 10.1093/sysbio/syy032
- Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302. doi: 10.1093/molbev/msx248
- Ruhlman, T. A., and Jansen, R. K. (2014). “The plastid genomes of flowering plants,” in *Methods in molecular biology* (Springer, New York, NY), 3–38. doi: 10.1007/978-1-62703-995-6_1
- Shen, X., Wu, M., Liao, B., Liu, Z., Bai, R., Xiao, S., et al. (2017). Complete chloroplast genome sequence and phylogenetic analysis of the medicinal plant *Artemisia annua*. *Molecules* 22, 1330. doi: 10.3390/molecules22081330
- Shi, L., Chen, H., Jiang, M., Wang, L., Wu, X., Huang, L., et al. (2019). CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47, W65–W73. doi: 10.1093/nar/gkz345
- Shi, W., Song, W., Liu, J., Shi, C., and Wang, S. (2023). Comparative chloroplast genome analysis of *Citrus* (Rutaceae) species: Insights into genomic characterization, phylogenetic relationships, and discrimination of subgenera. *Sci. Hortic.* 313, 111909. doi: 10.1016/j.scienta.2023.111909
- Sorensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Biologiske skrifter* 5, 1–34.
- Straub, S. C. K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R. C., and Liston, A. (2012). Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *Am. J. Bot.* 99, 349–364. doi: 10.3732/ajb.1100335

- Taha, A., Ettaqy, A., Mderssa, M. E., Belaqqiz, M., Fokar, M., Boukcm, H., et al. (2023). Comprehensive review of morphological adaptations and conservation strategies of cactiform succulents: A case study of *Euphorbia* species in arid ecosystems. *Biosyst. Divers.* 31, 358–367. doi: 10.15421/012342
- Taha, A., Zoubi, A., Ettaqy, A., El-Mderssa, M., Belaqqiz, M., Fokar, M., et al. (2024). Environmental drivers of *Euphorbia resinifera* seed germination and seedling establishment for conservation purpose. *Braz. J. Biol.* 84, e281196. doi: 10.1590/1519-6984.281196
- Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E. S., Fischer, A., Bock, R., et al. (2017). GeSeq - Versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* 45, W6–W11. doi: 10.1093/nar/gkx391
- Vieira, L. N., Dos Anjos, K. G., Faoro, H., Fraga, H. P. F., Greco, T. M., Pedrosa, F. O., et al. (2016). Phylogenetic inference and SSR characterization of tropical woody bamboos tribe *Bambuseae* (Poaceae: Bambusoideae) based on complete plastid genome sequences. *Curr. Genet.* 62, 443–453. doi: 10.1007/s00294-015-0549-z
- Wang, A. H. J., Hakoshima, T., van der Marel, G., van Boom, J. H., and Rich, A. (1984). AT base pairs are less stable than GC base pairs in Z-DNA: the crystal structure of d (m5CGTAm5CG). *Cell* 37, 321–331. doi: 10.1016/0092-8674(84)90328-3
- Wang, T., Li, X., Tang, C., Cao, Z., He, H., Ma, X., et al. (2024). Complete chloroplast genomes and phylogenetic relationships of *Pedicularis chinensis* and *Pedicularis kansuensis*. *Sci. Rep.* 14, 13548. doi: 10.1038/s41598-024-63815-0
- Wang, Y., Wang, S., Liu, Y., Yuan, Q., Sun, J., and Guo, L. (2021). Chloroplast genome variation and phylogenetic relationships of *Atractylodes* species. *BMC Genomics* 22, 103. doi: 10.1186/s12864-021-07394-8
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58, 236–244. doi: 10.1080/01621459.1963.10500845
- Wei, N., Pérez-Escobar, O. A., Musili, P. M., Huang, W. C., Yang, J. B., Hu, A. Q., et al. (2021). Plastome evolution in the hyperdiverse genus *Euphorbia* (Euphorbiaceae) using phylogenomic and comparative analyses: Large-scale expansion and contraction of the inverted repeat region. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.712064
- Wheeler, G. L., Dorman, H. E., Buchanan, A., Challagundla, L., and Wallace, L. E. (2014). A review of the prevalence, utility, and caveats of using chloroplast simple sequence repeats for studies of plant biology. *Appl. Plant Sci.* 2, 1400059. doi: 10.3732/apps.1400059
- Wilson, T., and Munro, N. (2019). *Euphorbia makallensis* Carter, a northern Ethiopian of very limited distribution. *Br. Cactus Succul. Soc.* 37, 43–46.
- Wu, Y., Liu, F., Yang, D.-G., Li, W., Zhou, X.-J., Pei, X.-Y., et al. (2018). Comparative chloroplast genomics of *Gossypium* species: insights into repeat sequence variations and phylogeny. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00376
- Xu, J., Zhang, H., Yang, F., Zhu, W., Li, Q., Cao, Z., et al. (2025). Phylogeny of *Camphora* and *Cinnamomum* (Lauraceae) based on plastome and nuclear ribosomal DNA data. *Int. J. Mol. Sci.* 26, 1370. doi: 10.3390/ijms26031370
- Zhao, S., Gao, X., Yu, X., Yuan, T., Zhang, G., Liu, C., et al. (2024). Comparative analysis of chloroplast genome of *Meconopsis* (Papaveraceae) provides insights into their genomic evolution and adaptation to high elevation. *Int. J. Mol. Sci.* 25, 2193. doi: 10.3390/ijms25042193
- Zhu, M., Feng, P., Ping, J., Li, J., Su, Y., and Wang, T. (2021). Phylogenetic significance of the characteristics of simple sequence repeats at the genus level based on the complete chloroplast genome sequences of Cyatheaceae. *Ecol. Evol.* 11, 14327–14340. doi: 10.1002/ece3.8151
- Zhu, A., Guo, W., Gupta, S., Fan, W., and Mower, J. P. (2016). Evolutionary dynamics of the plastid inverted repeat: The effects of expansion, contraction, and loss on substitution rates. *New Phytol.* 209, 1747–1756. doi: 10.1111/nph.13743