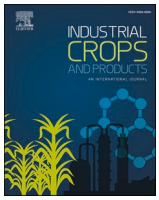
Industrial Crops & Products 234 (2025) 121477

Contents lists available at ScienceDirect

Industrial Crops & Products

journal homepage: www.elsevier.com/locate/indcrop



Comprehensive plastome analysis in prickly pear cactus (*Opuntia*, Cactaceae), a key genus for sustainable fuel and food: Uncovering evolutionary dynamics and germplasm

Cheng Chen [a,c,1], Jizhe Han [a,b,1], Xinyu Bai [a,b], Jing Yan [a], Yuqing Feng [a,b], Jing Liu [a,b], Xinyan Xu [a,b], Yinan Xiang [a,b], Chen Lin [a], Huiru Li [a], Yonghui Liu [a,b], Jing Yu [c], Yangchun Han [d], Zhechen Qi [b,*], Xiaoling Yan [a,*]

[a] *Eastern China Conservation Centre of Wild Endangered Plant Resources, CAS Center for Excellence in Molecular Plant Sciences Chenshan Plant Science Research Center, Shanghai Chenshan Botanical Garden, Shanghai 201602, China*
[b] *Zhejiang Province Key Laboratory of Plant Secondary Metabolism and Regulation, College of Life Sciences and Medicine, Zhejiang Sci-Tech University, Hangzhou 310018, China*
[c] *College of Life Science, Shanghai Normal University, Shanghai 200234, China*
[d] *Integrated Technical Service Center of Jiangyin Customs, Jiangyin 214400, China*

ARTICLE INFO

Keywords:
Opuntia
Inverted repeat variation
Ancestral state reconstruction
Divergence time estimation
Plastid SSR
Germplasm identification
Chloroplast genome evolution

ABSTRACT

Opuntia, a genus in Cactaceae, holds considerable economic significance due to its diverse uses. However, species identification and understanding of phylogenetic relationships within this genus are challenging, given the extensive morphological variation and frequent hybridization. This study analyzed plastomes from 40 *Opuntia* and 32 additional species within Cactaceae, focusing on plastome structure, gene content, and evolutionary dynamics. Our findings reveal significant structural variations and changes in gene content across *Opuntia* plastomes, with sizes ranging from 121,985 bp to 152,717 bp. These variations are primarily driven by the contraction and expansion of Inverted Repeat (IR) regions, which range from 338 to 30,530 bp. We identified 101 species-specific SSRs across the *Opuntia* plastomes, along with six high-resolution plastid markers (*ndhJ-trnF*, *atpE-trnM-CAU*, *psaJ-rpl33*, *ndhC-rbcL*, *psaA-ycf3*, *atpB-rbcL*), which offer potential tools for accurate germplasm identification. Our phylogenomic framework within Cactaceae traced the ancestral IR states of *Opuntia* through divergence time and ancestral trait reconstruction, which suggests an origin in the Late Miocene (approximately 8.54 million years ago). Initial slight IR contraction was observed, followed by diversification involving extreme contraction and expansion events. These events, potentially adaptive responses to climate changes, are proposed based on patterns of IR variation and documented climate shifts during the Late Tertiary and Quaternary periods These findings provide key molecular markers that could help resolve taxonomic challenges within the *Opuntia* genus and enhance our understanding of *Opuntia* plastome evolution. This study offers comprehensive genomic resources for this economically significant genus, laying a foundation for future species identification and breeding applications.

1. Introduction

Opuntia, comprising approximately 154 species of succulent plants (Korotkova et al., 2021), thrives predominantly in the arid and semiarid regions of the Americas and is cultivated globally (Anderson, 2001). Commonly known as prickly pears, these species hold significant value in agriculture, medicine, and various industries (Yahia and Saenz, 2011; Chahdoura et al., 2016; Ciriminna et al., 2019; Kumar and Sharma, 2020; Ramadan et al., 2021). Rich in polyphenols, flavonoids, and other bioactive compounds, *Opuntia* exhibits pharmacological properties including wound healing, antibacterial, antioxidant, and anti-inflammatory effects (Koch and Kennedy, 1980; Konings, 2010; Ammar et al., 2015; Santos-Díaz et al., 2017; Aruwa et al., 2018). Its therapeutic potential also extends to managing chronic diseases, such as

* Corresponding authors.
E-mail addresses: zqi@zstu.edu.cn (Z. Qi), xlyan@cemps.ac.cn (X. Yan).
[1] These authors contributed equally to the work.

https://doi.org/10.1016/j.indcrop.2025.121477
Received 3 January 2024; Received in revised form 11 April 2025; Accepted 4 July 2025
Available online 15 July 2025

diabetes, obesity, and cardiovascular disorders (Ammar et al., 2015; Chahdoura et al., 2015, 2016; Santos-Díaz et al., 2017; Kumar and Sharma, 2020; Ramadan et al., 2021; Zeghbib et al., 2022; Carpena et al., 2023). Moreover, *Opuntia* has expanding applications in biotechnology, food processing, and material sciences, further underscoring its industrial significance (Barba et al., 2017; Quintanar-Orozco et al., 2018; Ramadan et al., 2021; Carpena et al., 2023).

Beyond its industrial applications, *Opuntia* aligns well with environmental sustainability and bioenergy initiatives. Its high water-use-efficiency, drought resilience, and rapid growth, make it a strong candidate for bioenergy production, particularly in arid regions (Koch and Kennedy, 1980; Yahia and Saenz, 2011; Mason et al., 2015; Yang et al., 2019; Kumar and Sharma, 2020; Jorge et al., 2023). Recognized by the Food and Agriculture Organization (FAO) as a strategic crop, *Opuntia*, especially *O. ficus-indica*, is widely cultivated for food, feed, and industrial purposes, contributing to a global market valued at $1.9 billion in 2020 (Ramadan et al., 2021; Lu et al., 2023; Rodrigues et al., 2023; Shoukat et al., 2023).

Despite its economic significance, *Opuntia* presents challenges in species identification due to high morphological similarity, intraspecific variation, hybridization, synonymic complexity, and polyploidy (Labra et al., 2003; Majure et al., 2012; Valadez-Moctezuma et al., 2022). Accurate identification is crucial for effective germplasm utilization and evolutionary studies (Griffith, 2004; Las Casas et al., 2018; Challa and Neelapu, 2019; Mao et al., 2021; Rache-Cardenal et al., 2022). Current taxonomy primarily relies on a combination of vegetative and reproductive traits, such as growth form, cladode and spine characteristics, number of areoles, as well as fruit and floral morphology (Pinkava, 2003; Majure et al., 2017; Galicia-Pérez et al., 2023). However, similarities in these traits often complicate classification, highlighting the necessity for more effective genetic tools.

Molecular markers have improved phylogenetic resolution but remain limited in minor clades, leading to taxonomic inconsistencies (Nyffeler, 2002; Majure et al., 2012; Granados-Aguilar et al., 2022). Compared to nuclear markers, plastid genomes/locus offer advantages such as uniparental inheritance, lower recombination rates, and structural conservation, making them particularly valuable for phylogenetic studies in complex groups like *Opuntia*. Recent identification of ten informative plastid markers in Opuntioideae has significantly clarified intergeneric and tribal relationships (Köhler et al., 2020, 2023). Developing high-resolution plastid markers specifically for *Opuntia* will further enhance taxonomic clarity and provide deeper insights into its evolutionary history.

Plastid genomes typically contain a standard quadripartite circular structure with one long single copy (LSC), one short single copy (SSC) and two inverted repeat (IR) regions. However, an extensive analysis of 2511 plastomes across diverse lineages revealed that approximately 10.31 % of species exhibit loss of IR regions (Mohanta et al., 2020), a phenomenon noted in families such as Euphorbiaceae, Fabaceae, Passifloraceae, and Cactaceae (Sanderson et al., 2015; Cauz-Santos et al., 2020; Lee et al., 2021; Wei et al., 2021). This IR variation is thought to result from repeat-mediated intra- and inter-molecular recombination, impacting the stability and evolutionary trajectory of plastomes (Cauz-Santos et al., 2020; Lee et al., 2021; Wei et al., 2021).

Within Cactaceae, substantial variations in plastome gene content, order, and structure of plastomes have been linked to changes in the length and composition of the IR regions (Solórzano et al., 2019; Köhler et al., 2020; Dalla Costa et al., 2022; Yu et al., 2023; Köhler et al., 2023). In *Opuntia*, Yu et al. (2023) identified a significant IR contraction in *O. microdasys*, much shorter than the typical angiosperm plastome of ~25 kb. Köhler et al. (2023) further recognized six distinct types of plastomes in *Opuntia* based on IR patterns, with lengths ranging from 657 bp to 32,085 bp. These dynamic changes in IRs are thought to be associated with the plastid DNA repair process, potentially reflecting adaptations to arid and semi-arid environments (Sanderson et al., 2015; Köhler et al., 2020, 2023; Yu et al., 2023). However, no studies have yet explored these IR variations from an evolutionary perspective, particularly regarding their potential role in species' evolutionary processes or the existence of broader evolutionary patterns. Research on this topic, both temporally and in terms of trait reconstruction, remains lacking.

In this study, we embarked on a comprehensive analysis of the plastomes from 40 *Opuntia* species, along with 32 additional species within Cactaceae and two outgroup species. Our research aimed to expand the existing plastome database for the *Opuntia* genus, enhance understanding of its phylogenetic relationships, investigate the structure of *Opuntia* plastomes, identify potential variations, and trace the ancestral IR state and its evolutionary trajectory.

2. Materials and methods

2.1. Samples collection, DNA extraction, and sequencing

For this study, nine *Opuntia* species were collected (accessions are provided in Table B1) and sequenced using the genomic skimming approach. Additionally, plastomes of 31 *Opuntia* species were obtained from the NCBI database (Table B2), resulting in a total of 40 *Opuntia* species for analysis. To study the plastome structural evolution within the Cactaceae family, we acquired 34 plastome sequences from NCBI and Dryad (Table B2). This expanded dataset encompasses 74 species representing eight well-supported tribes across three Cactaceae subfamilies: Pereskioideae, Opuntioideae, and Cactoideae. Two outgroup species, *Portulaca oleraceae* (Portulacaceae) and *Talinum paniculatum* (Talinaceae), were also include. To optimize DNA extraction, we prepared dry cactus stem epidermis to reduce excess mucilage. The extraction of total genomic DNA was conducted using the TIANGEN Plant Genomic DNA kit (Beijing, China), following the manufacturer's instructions. DNA concentration and quality were assessed with a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA) and verified through 1 % agarose gel electrophoresis. Only DNA samples that met our quality standards was sent to Majorbio Co., Ltd (Shanghai, China) for library preparation and high-throughput sequencing. Sequencing was performed on the Illumina NovaSeq 6000 Platform, generating 150 bp paired-end reads.

2.2. Plastome De Novo assembly and annotation

Initial quality control and adaptor removal from the DNA pair-end raw reads were performed using Trimmomatic v0.32 (Bolger et al., 2014). Subsequently, the clean reads were then aligned to the complete plastome of *Opuntia quimilo* (Köhler et al., 2020) using BWA v0.7.17 (Li, 2013) to evaluate sequence quality and depth. The *de novo* assembly of the plastomes was achieved from these clean reads using GetOrganelle v1.7.7.0 (Jin et al., 2020) following the recommended parameters, which integrates Bowtie2 (Langmead and Salzberg, 2012), BLAST (Camacho et al., 2009) and SPAdes 3.1.0 (Bankevich et al., 2012). The assembled graph files were imported into Bandage (Wick et al., 2015) for visualization and checking assembly result and quadripartite structure (LSC, SSC and two IRs). Qualified plastid genomes were annotated (genes, coding regions, RNA and IR regions) using PGA (Qu et al., 2019) with published plastid genome (*Opuntia quimilo*, GenBank accession: NC083949) as reference and refined through meticulous manual inspection and adjustments in Geneious v9.0.2 (Biomatters Ltd., Auckland, New Zealand, https://www.geneious.com/). Missing genes from the PGA annotation were manually annotated using the "Map to Reference(s)" function, with the closest related species serving as a reference. All structure variation and genes transcription were manually checked in Geneious and found a putative pseudogene *accD*.

2.3. Whole plastome comparative analysis and codon usage

The basic statistics of the plastid genome (such as length, GC content, and gene count, and number of IR regions) were analyzed using

Geneious v9.0.2, and each gene sequence was carefully examined to identify pseudogenes resulting from premature stop codons or disruption at the quadripartite boundary. To identify highly variable regions within *Opuntia*, the mVISTA program (Brudno et al., 2003) in Shuffle-LAGAN mode was employed, using *O. arechavaletae* (with one IR removed) as the reference for plastome sequence comparison. The progressive mauve algorithm (Darling et al., 2010) facilitated comparative analysis of locally collinear blocks (LCBs) and genomic rearrangements among the plastomes. Nucleotide polymorphism (*Pi*) within *Opuntia* was quantified using DnaSP v6 (Rozas et al., 2017) for both coding and non-coding regions, with visualization achieved through the R package "ggplot2" (Wickham, 2011). Identification markers were selected based on regions exhibiting high *Pi* values and significant divergence, coupled with appropriate sequence lengths. To assess the resolution and stability of these markers among *Opuntia* species, we constructed a Neighbor-Joining (NJ) tree using Tree Builder in Geneious v9.0.2. This approach allowed us to evaluate the effectiveness of the selected markers in differentiating between species within the *Opuntia* genus. Additionally, primer pairs for amplification of these markers were designed using Primer3 v2.3.7 (Untergasser et al., 2012). The comparative lengths and IR regions boundaries of the quadripartite structures in the plastomes were illustrated using CPJSdraw (Li et al., 2023) and also in progressive mauve algorithm and in a handmade chloroplast genome gene order plot. The MIcroSAtellite Identification Tool (MISA) (Beier et al., 2017) was employed to identify the Simple Sequence Repeats (SSR) within the *Opuntia* plastomes. Additionally, we analyzed the nonsynonymous and synonymous substitution rates among *Opuntia* chloroplast genes using the KaKs_calculator (Wang et al., 2010). Relative Synonymous Codon Usage (RSCU) and Effective Number of Codon (ENC) analysis were conducted among the protein-coding sequences of plastomes within *Opuntia* to provide insights into the codon usage patterns, using CodonW 1.4.4 (John, 2004).

2.4. Phylogenetic analysis

Considering the notable structural rearrangements observed in the plastomes of the Cactaceae family, our phylogenetic analysis primarily focused on protein-coding sequences (CDS) of the entire plastome. The analysis encompassed 72 representative species from 24 genera spanning three Cactaceae subfamilies, in addition to two outgroup species. The CDS sequences were meticulously extracted from the genebank files and concatenated using two python scripts (https://github.com/Kinggerm/PersonalUtilities). These sequences were then aligned using MAFFT v7.505 (Katoh et al., 2002). For the construction of the maximum likelihood (ML) phylogenetic tree, IQ-TREE v1.6.8 (Minh et al., 2020) was employed with specific parameters: -m MFP -bb 1000 -bnni. The resulting phylogenetic tree was annotated using the R package ggtree (Yu et al., 2017).

2.5. Divergence time estimation

The divergence time analysis was conducted using the CDS matrix of 74 plastomes, employing the BEAST v2.5 (Bouckaert et al., 2019). We implemented a relaxed lognormal molecular clock and a birth-death model. The best nucleotide model, GTR, was determined by ModelFinder (Kalyaanamoorthy et al., 2017). The tree topology was set as constraint using the ML phylogram derived from our phylogenetic analysis. Due to the absence of reliable Cactaceae fossils, we utilized estimated divergence times extrapolated from a comprehensive chronogram encompassing the broader phylogenetic framework of Caryophyllales plastome evolution (Yao et al., 2019). We established two calibration points with normal prior distributions in our dataset: the mean root divergence time at 67.15 million years ago (Ma) with a standard deviation (sigma) of 4.48, and the mean crown group age of the Opuntieae at 17.1 Ma with a sigma of 1.5. The Markov Chain Monte Carlo (MCMC) analysis was rigorously conducted for a substantial duration of one billion generations, ensuring adequate effective sample sizes for all parameters. Subsequently, the initial 25 % of samples were discarded as burn-in. Trees from the three independent runs were then merged to construct a maximum clade credibility (MCC) tree. This tree featured median branch lengths and 95 % highest posterior density (HPD) intervals at nodes, using TreeAnnotator (Bouckaert et al., 2019) within BEAST v2.5 package. The resulting MCC tree was visualized and edited using ChiPlot (Xie et al., 2023).

2.6. Ancestral state reconstruction

Although plastid IR length is a continuous quantitative trait, its evolutionary dynamics remain difficult to model due to complex underlying mechanisms. IR variation can occur gradually via stepwise boundary shifts and microhomology-mediated recombination, or rapidly through large-scale structural rearrangements such as unequal crossing-over or repeat-mediated recombination (Maréchal and Brisson, 2010; Zhu et al., 2016; Choi et al., 2019). Currently, no unified evolutionary model effectively captures both modes of change. To address this, we adopted a complementary analytical approach combining both continuous and discrete frameworks to reconstruct ancestral IR states. This dual approach allows for cross-validation and complements the uncertainty inherent in modeling traits with both gradual and punctuated evolutionary trajectories. For continuous trait analysis, we employed the "fastAnc" function in the R package phytools (Revell, 2012), which estimates ancestral states under a Brownian motion model. This method assumes a constant-rate random walk of trait evolution and was applied to a MCC tree generated from BEAST, using the observed IR length matrix. For discrete-state reconstruction, we categorized IR lengths into five biologically informed bins based on observed distribution patterns: 1) A: IR length 0–1000 bp; 2) B: IR length $>$ 1000 bp and $\leq$ 2500 bp; 3) C: IR length $>$ 2500 bp and $\leq$ 17,000 bp; 4) D: IR length $>$ 17,000 bp and $\leq$ 30,000 bp; 5) E: IR length $>$ 30,000 bp. Bayesian inference was conducted using BayesTraits v3.0.1 (Meade and Pagel, 2018), implemented via RASP v4.3 (Yu et al., 2015). Ancestral states and transition rates were estimated with MCMC sampling across 1000 random trees and one consensus tree. Model selection was conducted to identify the best-fitting model between ER (equal rates) and ARD (all rates different), based on Bayes factor comparisons. For this, one thousand random MCMC trees and one consensus tree from BEAST were imported into RASP along with an IR state matrix. Detailed analysis was then conducted using the 'Multiple States Reconstruction' feature in BayesTraits, accessible through the 'Reconstruction' module in RASP.

3. Results and discussion

3.1. Structural characteristics of Opuntia plastomes: length variation and gene loss

Our Illumina NovaSeq 6000 sequencing generated approximately 10 Gb of raw DNA paired-end reads per *Opuntia* sample, with clean reads ranging from 70,025,953 to 80,855,119 (Table B1). Mapping to the *O. quimilo* plastome yielded average depths from 2040× to 9381×, supporting robust *de novo* assembly of 10 plastomes (Fig. 1 & A1), including one from NCBI SRA data (*O. inaperta*, SRR12440019). Additionally, 30 publicly available plastomes were incorporated (Table B2). The chloroplast genome sizes ranged from 121,985 bp to 152,717 bp (Table 1), with structural variation largely attributable to IR contraction/expansion (see Section 3.2). The LSC, SSC, and IR regions ranged from 50,009 to 101,684 bp, 4104 to 47,708 bp, and 338 to 30,530 bp, respectively. Gene content varied accordingly, ranging from 110 to 139 genes per plastome, including 70–91 protein-coding genes, 30–38 tRNAs, and 4–8 rRNAs (Fig. 2, Table B3).

While most genes were present as single copies in each *Opuntia* plastome, several genes exhibited lineage-specific variation in copy number or presence/absence. These included five *ndh* genes (*ndhA*,

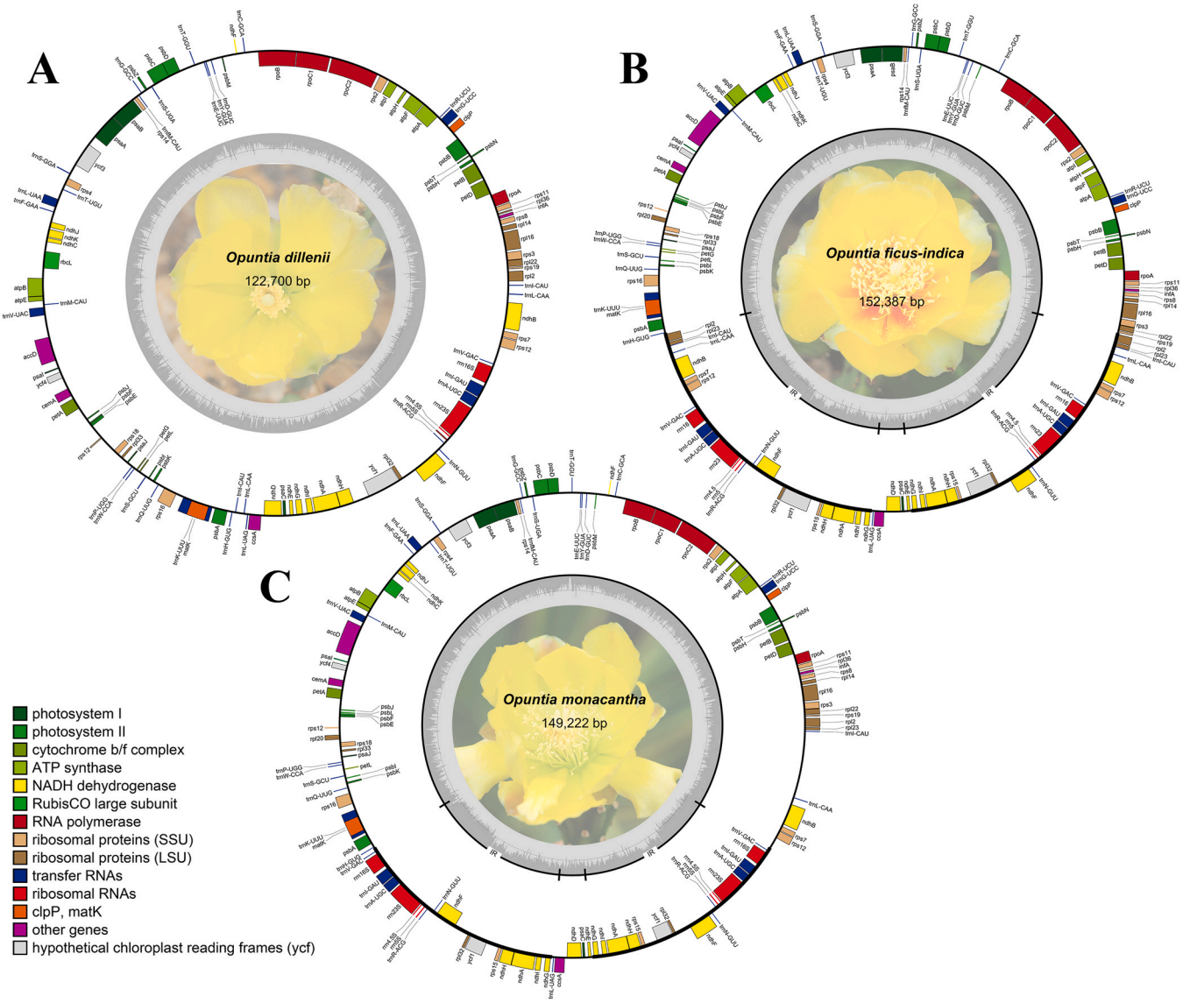


Fig. 1. Circular plastome maps of *Opuntia dillenii* (A), *O. ficus-indica* (B), and *O. monacantha* (C). Genes inside the circle are transcribed clockwise, and those outside, counter-clockwise. Black lines indicate the quadripartite structure: large single-copy (LSC), small single-copy (SSC), and the two inverted repeat regions (IRA and IRB). Gray shading represents GC content.

ndhB, *ndhF*, *ndhG*, *ndhH*), three *rpl* genes (*rpl2*, *rpl23*, *rpl32*), four *rps* genes (*rps7*, *rps12*, *rpl15*, *rpl19*), eight tRNA genes (*trnE-UUC*, *trnF-GAA*, *trnI-CAU*, *trnI-GAU*, *trnL-CAA*, *trnN-GUU*, *trnQ-UUG*, *trnV-GAC*), four rRNA genes (*rrn23*, *rrn16*, *rrn5*, *rrn4.5*), and two *ycf* genes (*ycf1*, *ycf2*) (Fig. 2). Notably, *ycf2* was entirely lost in 14 species; *ndhG* was absent in *Opuntia* sp. (GenBank accession: MW927506), while *ndhJ* and *ndhK* were missing in *O. polyacantha*. These losses likely occurred independently across different lineages. In addition, widespread pseudogenization of *accD*, *ycf1*, and *ycf2* was observed across the 40 plastomes, often accompanied by the loss of the *clpP* intron, suggesting coordinated structural streamlining. Beyond pseudogenization, *accD* also displayed substantial length variation among species. Nevertheless, its downstream region retained five conserved motifs (Fig. A2).

These frequent patterns of gene loss and pseudogenization in *Opuntia* plastomes—particularly involving *ndh*, *ycf*, and *accD* genes—likely reflect functional compensation rather than simple degenerative processes. While plastome reduction is often associated with holoparasitism (Chris Blazier et al., 2011; Lin et al., 2017; Kim et al., 2020), in *Opuntia*, the variation in copy number and partial loss of *ndh* genes may instead reflect adaptive responses to environmental stress, particularly under arid conditions (Sanderson et al., 2015; Cauz-Santos et al., 2020; Yu et al., 2023). This is consistent with the idea that gene loss does not necessarily compromise function, especially when alternative pathways (e.g., the antimycin A-sensitive cyclic electron flow) (Yamamoto et al., 2011; Labs et al., 2016) or compensatory mechanisms via nuclear or mitochondrial gene transfer are available (Lin et al., 2015; Yu et al., 2022)

The pseudogenization and extensive length variation of *accD* may similarly reflect ongoing functional relocation. Despite partial degradation, conserved downstream motifs—including core domains previously identified in Cereoideae (Yu et al., 2023), —suggest residual functionality or nuclear-encoded replacement (Rousseau-Gueutin et al., 2013). These gene losses and structural changes may reduce the energetic cost of plastid genome replication, particularly given the high copy number of plastomes per cell. By transferring dispensable or partially redundant genes to the nuclear genome, plants may enhance regulatory control and stress responsiveness (Forsythe et al., 2021; Kelly, 2021; Yu et al., 2023). This dynamic interaction between plastid and nuclear genomes reflects a broader co-evolutionary process in plants, in which endosymbiotic gene transfer contributes not only to genomic economy

Table 1
The summary of plastome characteristics in fourteen *Opuntia* species.

Species	*GenBank accession*	*Genome*						*LSC*		*SSC*		*IR*	
		Length	GC%	Gene	CDS	tRNA	rRNA	Length	GC%	Length	GC%	Length	GC%
Opuntia sp.	MW927506	122,617	36.00 %	115	78	33	4	90,059	33.81 %	31,882	37.88 %	338	37.20 %
Opuntia polyacantha	OQ613398	121,985	36.10 %	114	77	33	4	72,678	34.95 %	47,708	37.93 %	657	41.10 %
Opuntia dillenii	OR722820	122,700	36.07 %	114	78	32	4	88,469	35.21 %	32,067	38.48 %	1082	35.86 %
Opuntia rufida	NC083952	123,145	36.08 %	117	81	32	4	89,040	35.23 %	31,805	38.59 %	1150	36.26 %
Opuntia engelmannii var. cuija	OQ613386	123,275	36.10 %	114	78	32	4	88,784	35.27 %	32,085	38.39 %	1203	36.74 %
Opuntia setispina	NC083954	123,480	36.08 %	118	82	32	4	88,297	35.18 %	31,679	38.54 %	1752	36.07 %
Opuntia aureispina	OR722827	124,073	36.10 %	116	80	32	4	88,815	35.18 %	31,731	38.56 %	1763	36.30 %
Opuntia chisosensis	NC083934	123,987	36.10 %	118	82	32	4	88,729	35.20 %	31,732	38.56 %	1763	36.30 %
Opuntia macrocentra	NC083943	123,988	36.10 %	118	82	32	4	88,730	35.20 %	31,732	38.56 %	1763	36.30 %
Opuntia macrorhiza	NC083944	123,881	36.10 %	118	82	32	4	88,566	35.17 %	31,785	38.63 %	1765	35.98 %
Opuntia humifusa	OR722822	123,194	36.09 %	115	79	32	4	87,935	35.19 %	31,727	38.57 %	1766	36.24 %
Opuntia mesacantha	NC083945	123,784	36.10 %	118	82	32	4	88,535	35.23 %	31,717	38.56 %	1766	36.30 %
Opuntia austrina	NC083932	123,091	36.10 %	118	82	32	4	87,823	35.22 %	31,734	38.55 %	1767	36.28 %
Opuntia chlorotica	NC083935	123,748	36.10 %	118	82	32	4	88,511	35.23 %	31,703	38.56 %	1767	36.33 %
Opuntia drummondii	NC083938	123,611	36.10 %	118	82	32	4	88,346	35.26 %	31,731	38.56 %	1767	36.28 %
Opuntia strigil	NC083956	123,994	36.08 %	118	82	32	4	88,723	35.19 %	31,705	38.57 %	1783	36.23 %
Opuntia basilaris	OR722819	123,654	36.08 %	115	78	32	5	88,407	35.30 %	31,553	38.64 %	1847	33.03 %
Opuntia microdasys	NC083946	124,337	36.20 %	119	83	32	4	87,528	35.21 %	31,809	38.56 %	2500	37.92 %
Opuntia cochenillifera	NC087798	138,084	36.00 %	125	86	32	7	101,241	36.44 %	4107	33.14 %	16,368	34.80 %
Opuntia dejecta	NC083937	137,452	35.90 %	126	87	32	7	100,607	36.39 %	4107	33.14 %	16,368	34.80 %
Opuntia guatemalensis	NC083940	137,494	35.90 %	126	87	32	7	100,636	36.42 %	4104	33.19 %	16,377	34.79 %
Opuntia jamaicensis	NC083941	137,596	35.90 %	126	87	32	7	100,735	36.44 %	4107	33.11 %	16,377	34.79 %
Opuntia inaperta	SRR12440019	137,799	35.90 %	126	86	33	7	100,696	36.40 %	4107	33.11 %	16,498	34.67 %
Opuntia gaumeri	NC083939	137,818	35.90 %	126	87	32	7	100,711	36.41 %	4107	33.11 %	16,500	34.67 %
Opuntia caracassana	NC083933	138,141	35.90 %	126	87	32	7	101,028	36.43 %	4107	33.11 %	16,503	34.66 %
Opuntia auberi	NC083930	138,379	35.90 %	126	87	32	7	101,064	36.41 %	4107	33.14 %	16,604	34.57 %
Opuntia sulphurea	PV464448	147,442	36.70 %	128	85	35	8	101,498	36.40 %	4112	33.15 %	20,916	39.75 %
Opuntia retrorsa	NC083951	147,640	36.70 %	130	87	35	8	101,684	35.58 %	4114	33.13 %	20,921	39.75 %
Opuntia colubrina	NC083936	147,853	36.70 %	130	87	35	8	101,636	35.59 %	4115	33.07 %	21,051	39.62 %
Opuntia arechavaletae	NC083929	149,275	36.60 %	130	87	35	8	101,335	35.57 %	4108	33.06 %	21,916	39.34 %
Opuntia monacantha	OR722828	149,222	36.60 %	131	88	35	8	100,934	35.56 %	4110	32.97 %	22,089	39.18 %
Opuntia quimilo	NC083949	150,374	36.56 %	130	87	35	8	101,475	35.51 %	4115	32.98 %	22,392	39.56 %
Opuntia macbridei	NC083942	148,031	36.60 %	130	87	35	8	98,975	35.48 %	4104	33.19 %	22,476	39.38 %
Opuntia quitensis	NC083950	148,114	36.60 %	130	87	35	8	99,057	35.47 %	4105	33.18 %	22,476	39.39 %
Opuntia pilifera	OR722829	151,047	36.66 %	134	89	37	8	88,930	35.19 %	4107	33.02 %	29,005	39.17 %
Opuntia scheeri	NC083953	151,001	36.70 %	135	90	37	8	88,718	35.23 %	4105	33.03 %	29,089	39.19 %
Opuntia ficus-indica	OR722821	152,387	36.70 %	136	91	37	8	87,543	35.12 %	4106	33.02 %	30,369	39.23 %
Opuntia pycnantha	OR722830	152,541	36.67 %	135	90	37	8	87,604	35.13 %	4115	33.03 %	30,411	39.13 %
Opuntia pachyrrhiza	NC083947	152,690	36.60 %	138	93	37	8	87,641	35.09 %	4111	33.03 %	30,469	39.11 %
Opuntia stenopetala	NC083955	152,717	36.70 %	139	94	37	8	87,545	35.08 %	4112	33.03 %	30,530	39.16 %

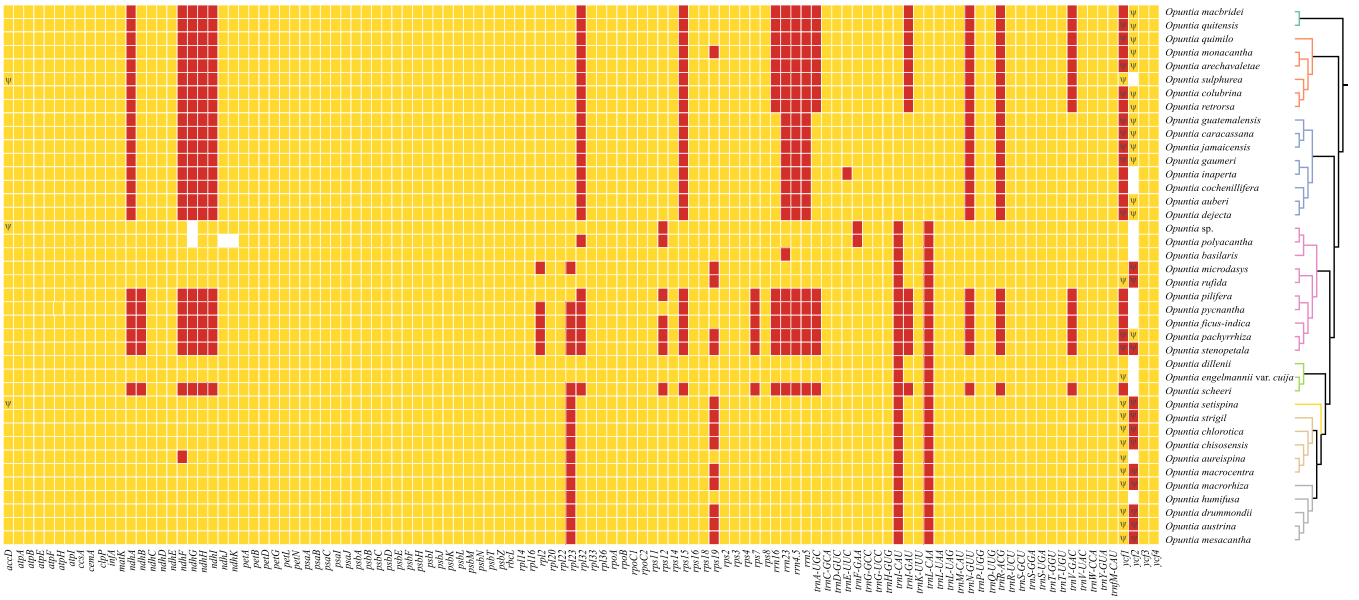


Fig. 2. Gene content comparison in *Opuntia* plastomes. Red blocks indicate two gene copies, yellow blocks a single copy, and white blocks denote gene absence. The right side illustrates the phylogenetic relationships within *Opuntia* based on phylogenomic relationships in Fig. 7.

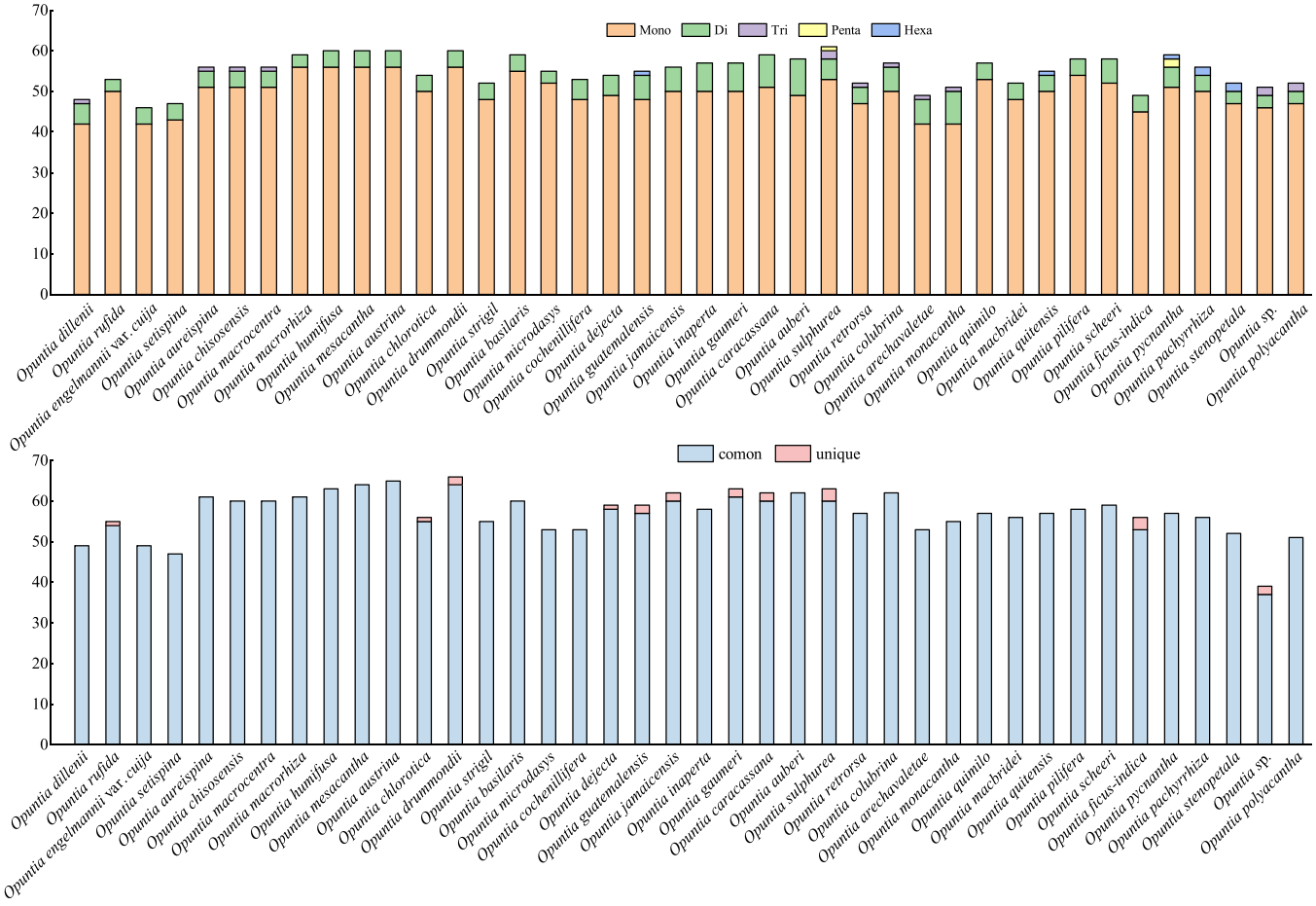


Fig. 3. Analysis of simple sequence repeats (SSRs) in *Opuntia* plastomes. (A) Number of unique and common SSRs detected in 40 *Opuntia* plastomes. (B) Frequency of identified SSRs types (Mono-, Di-. Tri-, Tetra, Penta- and Hexa- nucleotide repeats).

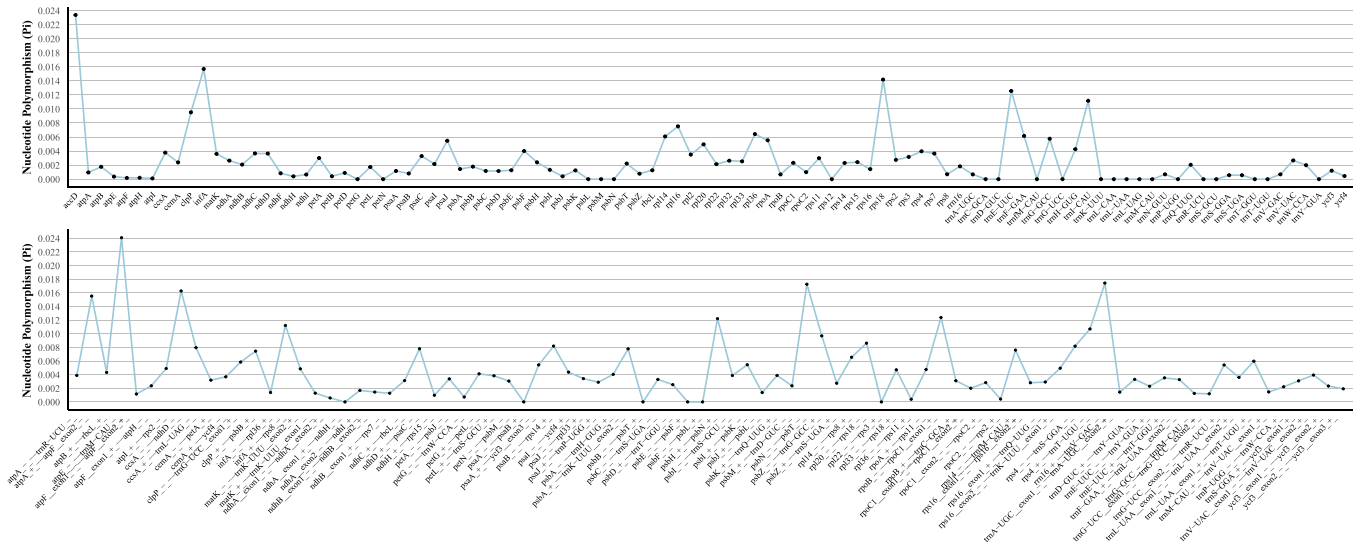


Fig. 4. Nucleotide polymorphism values (*Pi*) across *Opuntia* plastomes. Upper panel: *Pi* in coding regions; lower panel: *Pi* in non-coding regions. x-axis: coding genes or intergenic regions; y-axis: *Pi* value.

but also to functional innovation under selective pressures (Timmis et al., 2004; Weng et al., 2016; Kelly, 2021).

3.2. Potential markers of species identification: repeat elements, nucleotide diversity, and Ka/Ks ratios in Opuntia plastomes

We identified ptSSRs ranging from 51 in *O. engelmannii* var. *cuija* to 67 in *O. sulphurea*. The majority of ptSSRs in *Opuntia* plastomes are mono- (A/T, C/G) and di- nucleotides, with a presence of di-, tri-, penta-, and hexa- nucleotide repeats (Fig. 3, Table B4–B6). Twelve *Opuntia* species contained unique ptSSRs, varying from one in *O. strigil* to nine in *O sulphurea* (Fig. 3, Table B5), suggesting their potential as efficient species identification markers. Additionally, we selected a set of effective plastid markers based on nucleotide diversity (*Pi*) and comparative divergence analysis among *Opuntia* plastomes. The *Pi* values ranged from 0 to 0.02338 in coding regions and from 0 to 0.02408 in non-coding regions (Fig. 4 and Table B7). Regions with high *Pi* values were considered robust candidates for taxonomic markers. Furthermore, comparative analysis using mVISTA enabled the identification of highly divergent regions across the plastomes (Fig. A3). By integrating nucleotide diversity and genomic variation patterns and considering appropriate fragment lengths for practical use, we selected six promising non-coding markers: *ndhJ–trnF*, *atpE–trnM-CAU*, *psaJ–rpl33*, *ndhC–rbcL*, *psaA–ycf3*, *atpB–rbcL*). Except for *atpB–rbcL*, these regions have not been previously employed in phylogenetic studies of *Opuntia*. A NJ tree constructed using these six markers, including *Airamopoa erectoclada* and 40 *Opuntia* species, produced a topology (Fig. A4) largely consistent with that obtained using the whole CDS matrix (Fig. A5, see Section 3.4). Only minor topological differences were observed in the Basilaris, Macrocentra, and Humifusa clades. These markers exhibited strong phylogenetic resolution within *Opuntia*, with average bootstrap values exceeding 90.

Notably, Köhler et al. (2020) designed a separate set of ten plastid markers—including eight CDS (*accD*, *ycf1*, *ndhD*, *petD*, *ccsA*, *clpP*, *rpoC1*, *rpoC2*), one *trnK* intron, and one intergenic spacer (*psbE-rpl20*)—which were optimized for resolving phylogenetic relationships across tribes within the Opuntioideae subfamily. In contrast, none of their markers overlapped with the six intergenic regions identified in our study, which were specifically selected for their high discriminatory power at the species level within *Opuntia*. This contrast suggests that different taxonomic depths may require distinct, tailored marker sets. To facilitate future research and application, we also developed primer pairs for the six candidate intergenic regions (Table B5) as well as all species-specific ptSSRs identified in this study (Table B8). Integrating these newly identified markers with those previously used could provide a more comprehensive framework for resolving the phylogenetic and taxonomic complexities within *Opuntia*.

3.3. Selective pressure and codon usage bias in Opuntia plastomes

The Ka/Ks ratio analysis of *Opuntia* plastid genes showed that the majority of genes had Ka/Ks values less than 1 (Table B9), indicating overall patterns of purifying selection. However, a subset of genes—including the plastid-encoded RNA polymerase (PEP) subunits *rpoA*, *rpoB*, *rpoC1*, and *rpoC2*—exhibited Ka/Ks > 1 in multiple taxa, indicating strong signals of positive selection. When combined with phylogenetic clade information (Table B9; see Section 3.5 for detailed clade relationships), these patterns revealed lineage-specific variation in selective pressures acting on plastid-encoded genes across *Opuntia*. Notably, *rpoB* showed the most pronounced signal, with Ka/Ks values exceeding 2 in most species from Clades E to H, while *rpoA* exhibited elevated values specifically in Clade D. Within the same clade, increased Ka/Ks ratios were also observed for *atpA* and *rpl20*. Ribosomal protein genes *rps2*, *rps15*, and *rps18* displayed elevated values in particular lineages, with *rps2* exceeding 2 in most taxa of Clade B. In contrast, some genes lacked clear lineage-specific trends. For instance, *ndhB* showed Ka/Ks > 1 across a broad phylogenetic range, while *ycf1* and *clpP* exhibited sporadic elevations distributed across multiple clades.

Codon usage analysis across 40 *Opuntia* plastomes revealed a strong bias towards a set of 30 A/U, which were consistently preferred across species (Fig. A6). Codons such as UUA (Leu), AGA (Arg), GUU (Val), and AUU (Ile) exhibited RSCU values greater than 1.5 in nearly all taxa, while G/C-ending codons like CGC (Arg), CCC (Pro), and GGC (Gly) showed markedly lower RSCU values (< 0.6). The ENC values for these

Fig. 5. Plastome structure and gene order in *Opuntia* with different IR types (IRL, IRC, IRN, IRE). Linearized plastome maps for clarity. Dotted lines link identical genes or blocks across IR and LSC/SSC. Color coding indicates gene locations, with unique colors for each IR variation. Orange marks genes in IR region of *O. humifusa*. Red highlights genes in the IR region of *O. monacantha*. Sky blue and blue denote unique genes in the IR region of *O. ficus-indica*.

plastomes ranged from 33.62 to 59.42. In ENC-GC3 plots (Fig. A7), certain genes such as *rpl36*, *rbcL*, and *psbH* consistently deviated from the standard curve across multiple species.

The consistent signals of positive selection in different functional gene sets, particularly the PEP genes, suggest that adaptive evolution in the *Opuntia* plastid genome may have occurred in a clade-specific manner. Given the central role of the PEP complex in plastid transcription, the clade-specific patterns of selection observed may reflect evolutionary shifts in the composition or activity of the plastid transcriptional machinery (Vergara-Cruces et al., 2024; Wang et al., 2024). These shifts may have arisen in response to ecological diversification or heterogeneous environmental pressures experienced by different lineages (Piot et al., 2018; Gu et al., 2023; Wu et al., 2024). Further functional studies are needed to determine whether the observed amino acid changes influence the transcriptional dynamics of plastid genes or their interaction with nuclear-encoded sigma factors (Hwang et al., 2022; Zhang et al., 2023).

From another perspective, the codon usage patterns reflect the AT-rich nature of *Opuntia* plastid genomes and are highly conserved across lineages. Similar to the Ka/Ks < 1 values observed for most plastid genes, this suggests that both codon usage and substitution patterns may be shaped by shared evolutionary constraints, likely dominated by purifying selection (Qi et al., 2015; Fages-Lartaud et al., 2022). Such convergence in codon usage has also been reported in other plant lineages, showing strong intra-lineage consistency (Qi et al., 2015; Zhang et al., 2021), although intrageneric divergence has also been observed (Liu et al., 2014).

3.4. Structural variation in Opuntia plastomes: contraction and expansion of inverted repeat regions

Significant structural variations in the IR regions was observed among *Opuntia* species, with sizes ranging from 338 bp to an expansive 30,530 bp (Table 1, Fig. A8). Substantial IR expansion and contraction

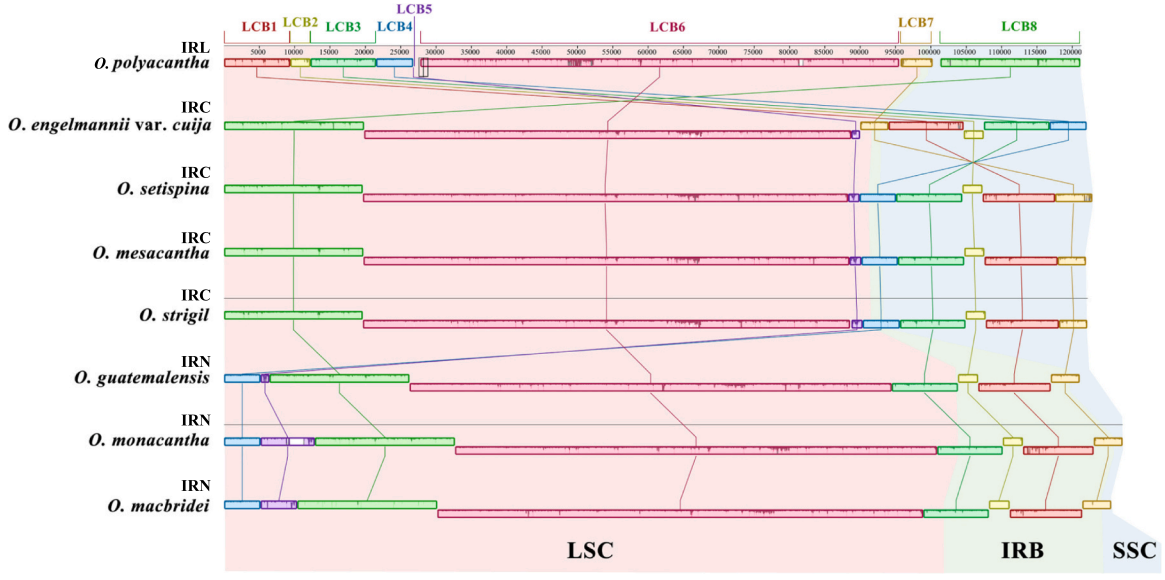


Fig. 6. Progressive mauve analysis of *Opuntia* plastomes using *O. dillenii* as reference (IRA removed). The figure illustrates the local collinear blocks (LCBs) across different *Opuntia* plastomes. Each colored blocks represent a distinct LCB, with their respective numbers and boundaries indicated. LCBs positioned below the horizontal black line represent regions of inversion. The color coding for different regions is as follows: pink for the large single-copy (LSC) region, green for the inverted repeat region B (IRB), and purple for the small single-copy (SSC) region.

led to marked genome size differences. Two species—*Opuntia* sp. (MW927506) and *O. polyacantha*—exhibited near-complete IR loss (<1 kb), while an additional. 15 species showed contracted IRs (1000–2500 bp; Table A7, Table B3). Together, species with IR loss or contraction accounted for approximately 45 % of the sampled taxa. In contrast, six species (~15 %), including *O. pilifera*, *O. ficus-indica*, and *O. stenopetala*, exhibited significantly expanded IRs (>29,000 bp). Based on IR length and gene content, four IR structural types were defined (Table 1, Table B10): 1) IRL: IR lost—nearly absent IRs (<1 kb, 0–1 gene); 2) IRC: IR contracted (<17 kb, 2–6 genes); 3) IRN: IR normal (16–23 kb, 12–17 genes); 4) IRE: IR expanded (>29 kb, 21–25 genes).

Marked differences in IR length, structure, and gene content were observed among representative plastomes of the four IR types (Fig. 5). In the IRL type, the IR region accounted for only ~1.5 % of the typical IRN length and was accompanied by an elongated SSC region that retained the *ndhI–trnV-GAC* gene block. The IRC type had slightly longer IRs (~7.9 % of IRN length), but instead of a reduced SSC, it gained several genes (e.g., *ndh*, *psaC*), while the LSC became relatively shorter. In the IRN plastome (IR ~22.4 kb), the SSC was greatly reduced to ~4.1 kb, and several genes formerly located in the SSC and IR regions were repositioned into the LSC. In the IRE type, IR expansion (~30.4 kb) did not notably enlarge the plastome (still ~150 kb). Instead of SSC contraction, parts of the LSC were incorporated into the IRs, duplicating genes like *rps7*, *ndhB*, and *ycf2*.

The expansion and contraction of IRs also affect regional GC content in *Opuntia* plastomes. While overall GC content varied slightly across the 40 plastomes (35.9–36.7 %; Table 1 & B3), regional differences were more pronounced. In species with contracted IRs, the IR regions had lower GC content (avg. ~36 %), about 3 % less than in IRN or IRE types. Notably, *O. basilaris* showed the lowest IR GC content at 33 %. Conversely, their SSC regions exhibited elevated GC content, averaging ~38.6 %—approximately 5.6 % higher than in species with IRN and IRE types.

Progressive Mauve analysis of representative species from each IR type identified eight local colinear blocks (LCBs) and revealed multiple structural rearrangements among *Opuntia* plastomes (Fig. 6). In the IRL-type species (*O. polyacantha*), LCBs 2, 5, and 6 were inverted, while LCBs 1–4 were relocated to the SSC in IRC plastomes. In *O. engelmannii* var. *cuija*, these blocks were reoriented, with LCB4 extending into the IR region in other IRC species. With IR expansion in IRN (*O. guatemalensis*) and IRE (*O. monacantha*, *O. macbridei*) types, LCB4 was transferred back into the LSC, together with LCB5. Despite structural rearrangements, the junctions between LSC/IR and SSC/IR were relatively conserved within the same IR type (Fig. A8). In IRC plastomes, *trnH* was consistently positioned near or at the JLB (LSC/IRb junction). In IRN plastomes, junction-associated genes remained consistent. In IRE types, *trnI* or *rpl23* flanked JLB and JLA, although their exact positions varied between species (e.g., *O. pilifera* vs. *O. scheeri*).

While IR variation has been reported in the subfamily Cereoideae of Cactaceae and in other angiosperm families such as Euphorbiaceae, Fabaceae, and Passifloraceae (Solórzano et al., 2019; Cauz-Santos et al., 2020; Mohanta et al., 2020; Wei et al., 2021; Yu et al., 2023), the concurrent presence of both extreme contraction and expansion within a single genus, as observed in *Opuntia*, remains highly uncommon. Although the exact role of IRs in plastids remains incompletely understood, most studies agree that IRs contribute significantly to structural stability, gene dosage balance, and functional conservation of the plastome (Maréchal et al., 2009; Wicke et al., 2011; Rabah et al., 2019; Mohanta et al., 2020). In *Opuntia*, the observed gene content differences, region-specific GC content shifts, and lineage-specific LCB rearrangements associated with varying IR sizes suggest that IRs may also play a pivotal role in plastome evolution (Köhler et al., 2023). These IR dynamics appear non-random and may reflect clade-specific genomic trajectories. Together, these findings provide a foundation for the phylogenetic and temporal investigation of IR evolution in *Opuntia*, which is further explored in Section 3.5.

3.5. Phylogenomic context and integrative evolutionary perspective on IR variation in Opuntia

The plastome phylogeny reconstructed from 72 protein-coding genes (57,668 bp) supports the monophyly of Opuntioideae and Cactoideae, with *Pereskia aculeata* recovered as the basal lineage (Fig. A5). Within *Opuntia*, three major clades were identified: two South American (SA) clades (Clades A and B) and a highly diversified North American (NA) clade, which is further divided into six subclades (Clades C–H). These eight clades correspond to the lineages previously described as Quitensis, Elatae, Nopalea, Basilares, Scheerianae, Setispina, Humifusa, and Macrocentra (Köhler et al., 2023; Majure et al., 2023). The accesssion *Opuntia* sp. (MW927506) from Chen et al. (2022), originally identified as *O. sulphurea*, was recovered within the Basilares clade (Clade D) alongside *O. basilaris* (Fig. A5). In contrast, *O. sulphurea* in our phylogeny was placed in the Elatae clade (Clade B). These results support the suggestion by Köhler et al. (2023) that the MW927506 accesssion may actually represent affinities of *O. polyacantha* rather than *O. sulphurea*.

Our time-calibrated phylogeny (Fig. 7, Table B11) places the origin of Cactaceae at approximately 61.89 Ma, with initial diversification around 44.49 Ma. Major divergence events within the subfamilies Cactoideae and Opuntioideae occurred at 27.08 Ma and 18.62 Ma, respectively. Within Opuntioideae, the crown ages of the tribes Cylindropuntieae, Tephrocacteae, and Opuntieae were estimated at 13.23 Ma, 15.42 Ma, and 14.77 Ma, respectively. The genus *Opuntia* originated around 10.42 Ma, with an initial split between SA-Clade A and SA-Clade B + NA clade at 8.54 Ma. A subsequent divergence separated SA-Clade B from the NA clade at approximately 8.12 Ma, followed by rapid diversification within the North American lineage during the late Miocene to early Quaternary (ca. 7.17–1.98 Ma), with most subclades forming within the past ~4 million years.

Both ancestral state reconstruction methods (Fig. 7 & A9) reveal a dynamic history of IR evolution across Cactaceae, with pronounced contraction and expansion events occurring mainly during the late Tertiary to Quaternary. Early in the family's evolutionary history (middle Eocene to early Miocene), IR length remained relatively stable (~9000 bp), consistent with the basal taxon *Pereskia aculeata*. While moderate IR variation was observed in most Cactoideae lineages, more drastic changes—including extreme contraction (e.g., *Mammillaria*, *Parodia*) and complete IR loss (*Carnegiea*)—emerged during the Late Miocene and Quaternary. In contrast, IR evolution within *Opuntia* displays a more complex and lineage-specific pattern. The ancestral IR length in early *Opuntia* (~8 Ma, Late Miocene) was estimated at approximately 16,000 bp, following a broader contraction observed in early Opuntioideae (~3000 bp around 17.9 Ma). In the early-diverging South American clades (Clades A & B), the IR expanded back to typical chloroplast sizes (~22,000 bp), suggesting a possible reversal of ancestral contraction during the Miocene-Pliocene transition. Meanwhile, the diversified North American clades exhibit a more heterogeneous pattern in IR length variation. The early-diverging NA Clade C retained an intermediate IR size (~16,000 bp), whereas the later-

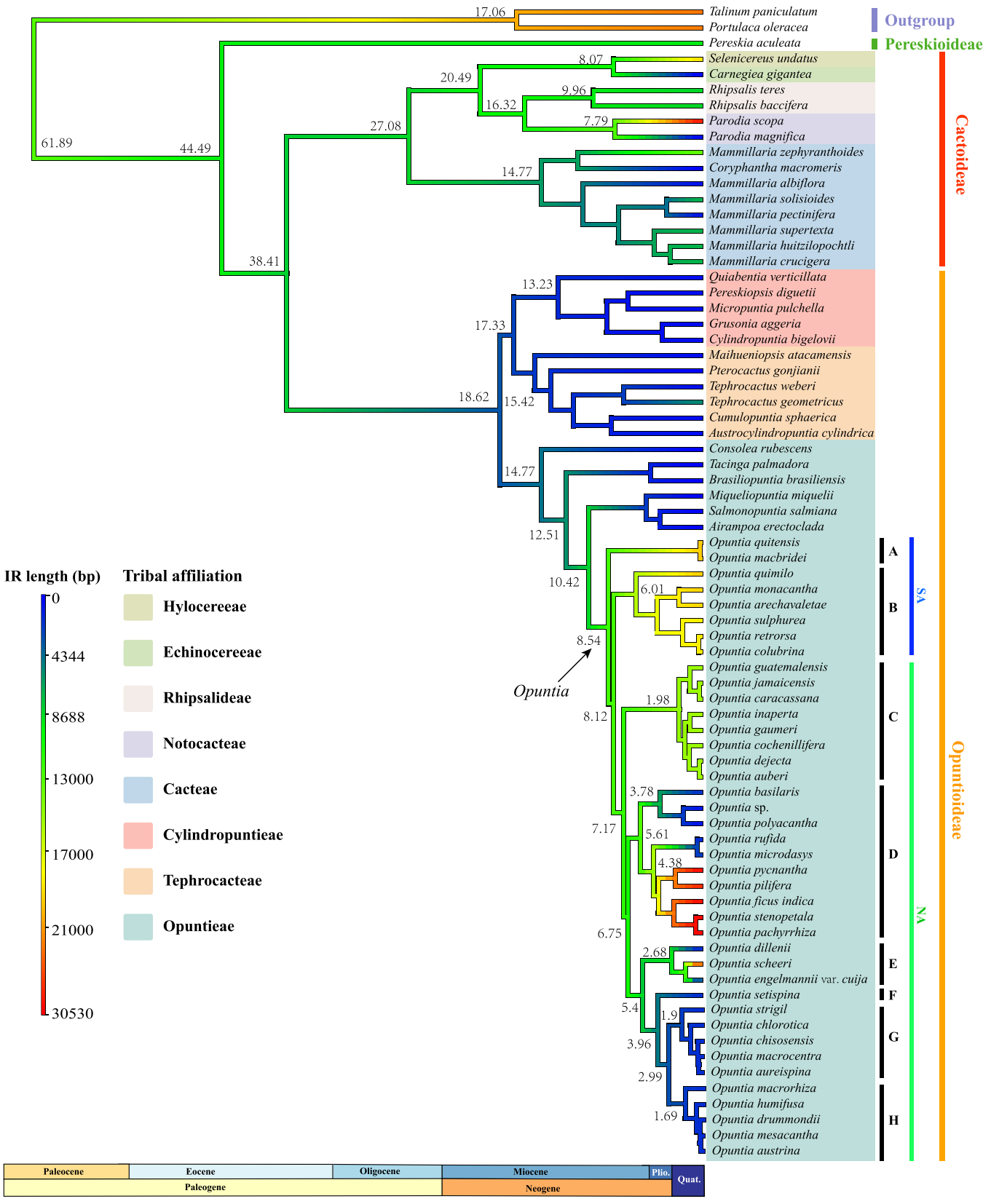


Fig. 7. Temporal divergence, phylogenetic relationships, and ancestral inverted repeat (IR) length state reconstruction in Cactaceae. Leaf label colors correspond to tribes within the Cactoideae and Opuntioideae subfamilies. The gradient bar, transitioning from dark blue to red, illustrates the reconstructed ancestral IR states across branches. Stripes on the right side of the figure distinguish the outgroup and the three Cactaceae subfamilies. Blue and green stripes behind *Opuntia* species denote their current distribution ranges in South America (SA) and North America (NA), respectively. Eight subclades (A, B, C, D, E, F, G H) were also highlighted. Numbers on nodes indicate divergence time in million of years ago (Ma). The timeline at the bottom provides a geological time scale measured in Ma.

diverging clades (D–H) exhibited more heterogeneous IR states. In Clade D, frequent and extreme IR contractions occurred early in its divergence (e.g., *O. polyacantha* and *Opuntia sp.*, representing the IRL type). However, during the Pliocene–Quaternary transition, striking IR expansions (>30,000 bp) were observed in several later-diverging lineages within Clade D. Among Clades E–H, persistent IR contraction was prevalent across most species—accounting for approximately 45 % of the sampled taxa—while IR expansion events also occurred in Clade E during the Quaternary (e.g., *O. scheeri*).

The inverted repeat (IR) regions are essential for plastome stability, mediating homologous recombination-based repair, buffering against deleterious structural rearrangements, and maintaining gene dosage balance through duplication of core genes (Maréchal et al., 2009; Wicke et al., 2011; Rabah et al., 2019; Mohanta et al., 2020). Despite their conserved role across most angiosperms, our results show that Opuntia plastomes exhibit both extreme IR contraction and expansion—especially within North American lineages during the Late Miocene to Quaternary (~7.17–1.98 Ma), a period of intensified aridification and climatic fluctuations (Velasco-de León et al., 2010; Sniderman et al., 2016; Bhattacharya et al., 2022).

The extensive variation in IR length, ranging from near-complete loss to expansions exceeding 30 kb, has direct consequences for gene copy number. IR expansion in species such as *O. pilifera* and *O. pycnantha* leads to the duplication of rRNA operons and other core genes (e.g., *rps7*, *ndhB*, *ycf2*), potentially increasing transcriptional output and buffering against environmental stressors. In this context, large IRs may enhance plastome robustness through increased recombinational repair and dosage stability under conditions of high temperature, intense UV radiation, and drought—characteristics of the American desert environments during the Miocene and Pleistocene (Comes and Kadereit, 1998; Hernández-Hernández et al., 2014).

However, plastid IRs are not indispensable. Multiple independent losses or contractions of IRs have been documented across plant lineages (Zhu et al., 2016; Choi et al., 2019; Mohanta et al., 2020; Wu et al., 2021). Recent studies show that plastomes lacking canonical IRs can still maintain gene expression balance via dose-compensatory mechanisms (Krämer et al., 2024). In *Opuntia*, lineages with contracted IRs may retain functional integrity despite structural reduction. Moreover, repeated IR loss or contraction may represent a resource-conserving strategy in metabolically constrained desert environments (Köhler et al., 2020). Accompanying IR-mediated gene loss or pseudogenization, EGT may occur, reducing plastome maintenance costs and reallocating resources—a process especially relevant under nutrient and water scarcity (Kelly, 2021; Yu et al., 2022; Majure et al., 2023).

In addition, reduced IR-mediated structural buffering may expose the plastome to greater plasticity. IR contraction has been associated with increased rearrangement frequency, including gene inversions, relocations, and deletions (Zhu et al., 2016). This flexibility may promote regulatory or metabolic innovation, particularly in taxa experiencing environmental instability (Wang et al., 2022).

The balance between plastome structural constraint and adaptive flexibility suggests that IR variation in *Opuntia* is not a neutral process, but reflects lineage-specific responses to environmental pressures, particularly those associated with aridification in the Late Miocene and Quaternary (Arakaki et al., 2011; Breslin et al., 2022; Majure et al., 2023). These dynamics may have co-evolved with key physiological adaptations, such as Crassulacean Acid Metabolism and succulent stem morphology, which jointly enhance drought resilience (Strand et al., 2019; Köhler et al., 2020, 2023). Similar extensive IR variation in Poales also suggest that such variation could be related to its successful diversification into diverse habitats (aquatic, terrestrial) and multiple photosynthetic pathway transitions (C3, C4) (Wu et al., 2024).

In addition to their evolutionary relevance, plastome-level changes—particularly IR dynamics—offer practical implications for germplasm conservation and breeding. Recent studies highlight the potential of using plastid genome variation to guide the selection of stress-resilient traits in crop improvement programs (Badawi et al., 2004; Li et al., 2021; Ahmad and Nixon, 2025). For *Opuntia*, a genus with high ecological and economic value, insights into plastome architecture may inform the development of varieties with enhanced drought tolerance and genomic stability under environmental stress. Therefore, understanding the evolutionary trajectory and functional consequences of IR variation not only sheds light on plastome plasticity and lineage diversification, but also provides a molecular basis for future *Opuntia* breeding strategies aimed at sustainable agriculture in changing climates (Comes and Kadereit, 1998; Davis et al., 2005; Hernández-Hernández et al., 2014).

4. Conclusion

This study provides a comprehensive assessment of plastome structural variation in *Opuntia*, revealing extensive and lineage-specific changes in IR regions, including both near-complete loss and extreme expansion. Through integrative phylogenomic and ancestral state analyses, we reconstruct the evolutionary dynamics of IR variation across clades and over time. In addition, the structural and sequence variation captured in this study offers valuable markers for improving species delimitation and accurate identification within *Opuntia*. These findings provide not only evolutionary perspectives but also practical genomic resources for species identification, germplasm management, and trait-oriented breeding in *Opuntia*.

CRediT authorship contribution statement

Cheng Chen: Writing – original draft, Visualization, Software, Methodology, Formal analysis, Data curation. **Jizhe Han:** Writing – review & editing, Software, Resources, Data curation. **Xinyu Bai:** Software, Methodology. **Jing Yan:** Validation, Resources, Methodology, Investigation. **Yuqing Feng:** Software, Resources. **Jing Liu:** Writing – review & editing, Visualization, Data curation. **Xinyan Xu:** Writing – original draft, Visualization, Methodology, Formal analysis, Data curation. **Yinan Xiang:** Writing – review & editing, Validation, Data curation. **Chen Lin:** Writing – review & editing, Resources. **Huiru Li:** Writing – review & editing, Resources. **Yonghui Liu:** Writing – review & editing, Resources. **Jing Yu:** Writing – review & editing, Validation. **Yangchun Han:** Writing – review & editing, Resources. **Zhechen Qi:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization. **Xiaoling Yan:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization.

Declaration of Competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study was financially supported by the Special Fund for Scientific Research of Shanghai Landscaping & City Appearance Administrative Bureau, grant numbers (G252409, G242412); the Natural Science Foundation of Zhejiang Province, grant number (LY21C030008); National Wild Plant Germplasm Resource Center, grant number (ZWGX2202).

Appendix A

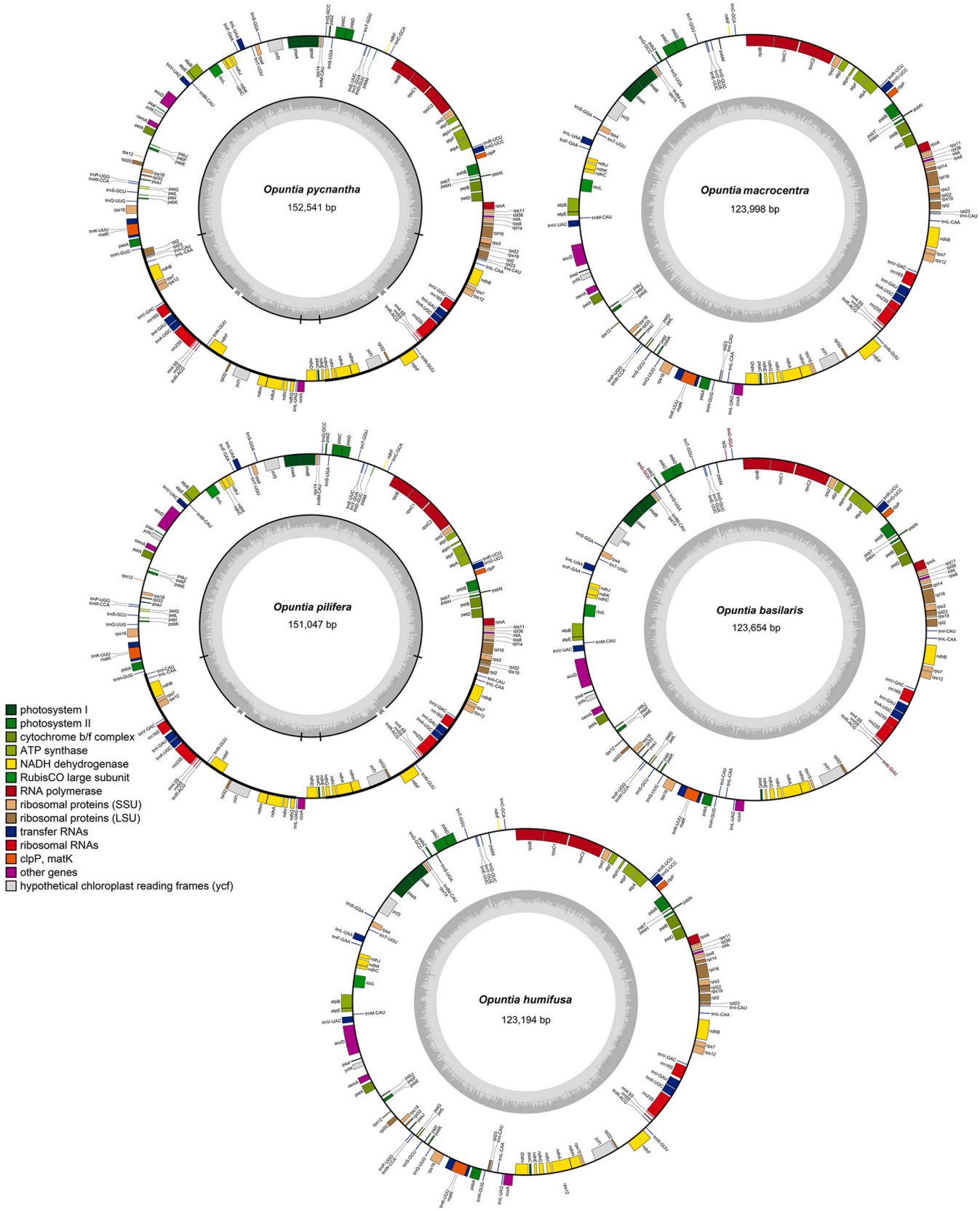


Fig. A1. Circular plastome maps of *Opuntia ficus-india*, *O. macrocentra*, *O. pilifera*, *O. basilaris*, and *O. humifusa*. Genes inside the circle are transcribed clockwise, and those outside, counter-clockwise. Black lines indicate the quadripartite structure: large single-copy (LSC), small single-copy (SSC), and the two inverted repeat regions (IRA and IRB). Gray shading represents GC content

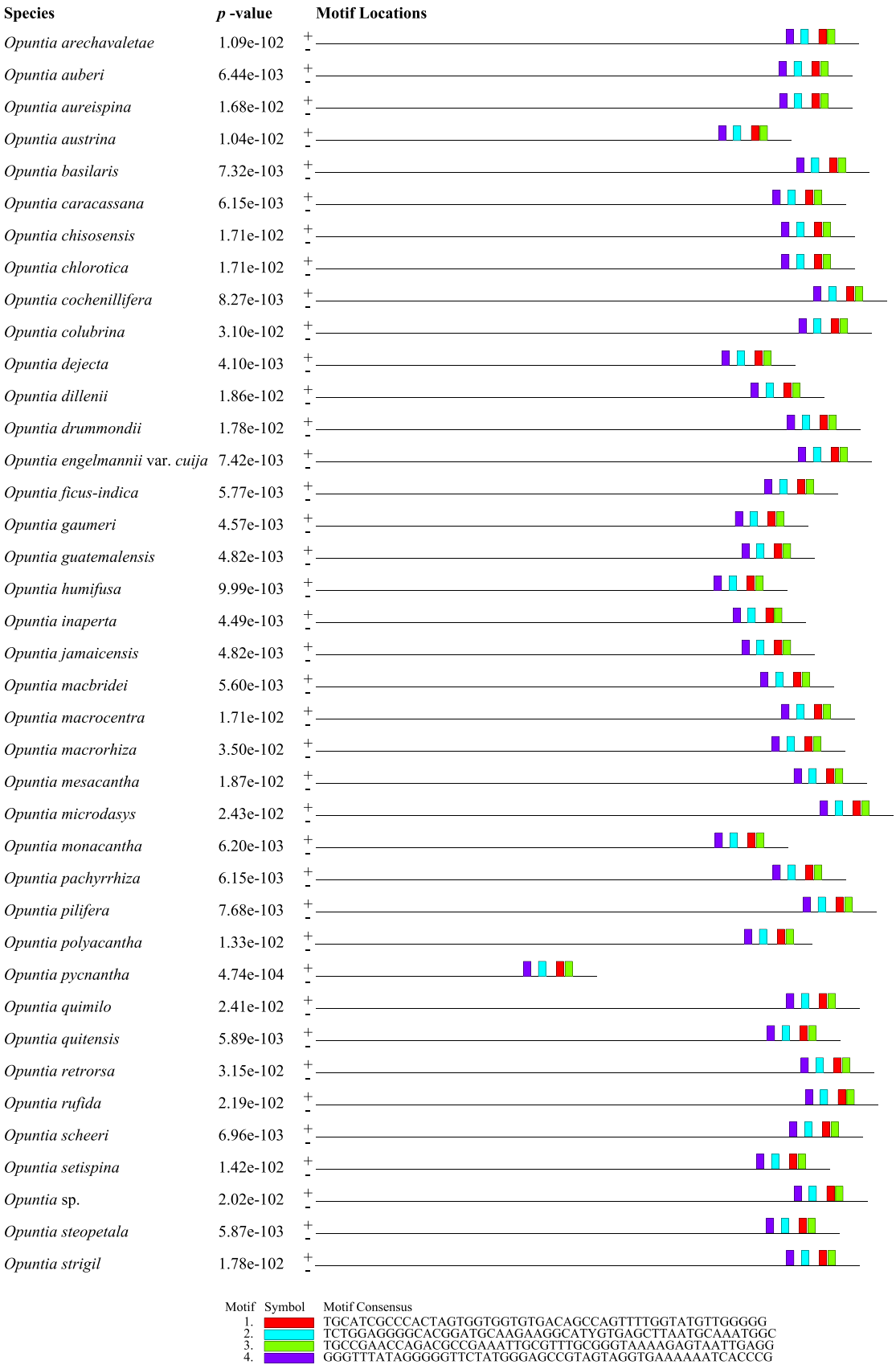


Fig. A2. The five conserved motifs of the *accD* gene across 39 *Opuntia* species

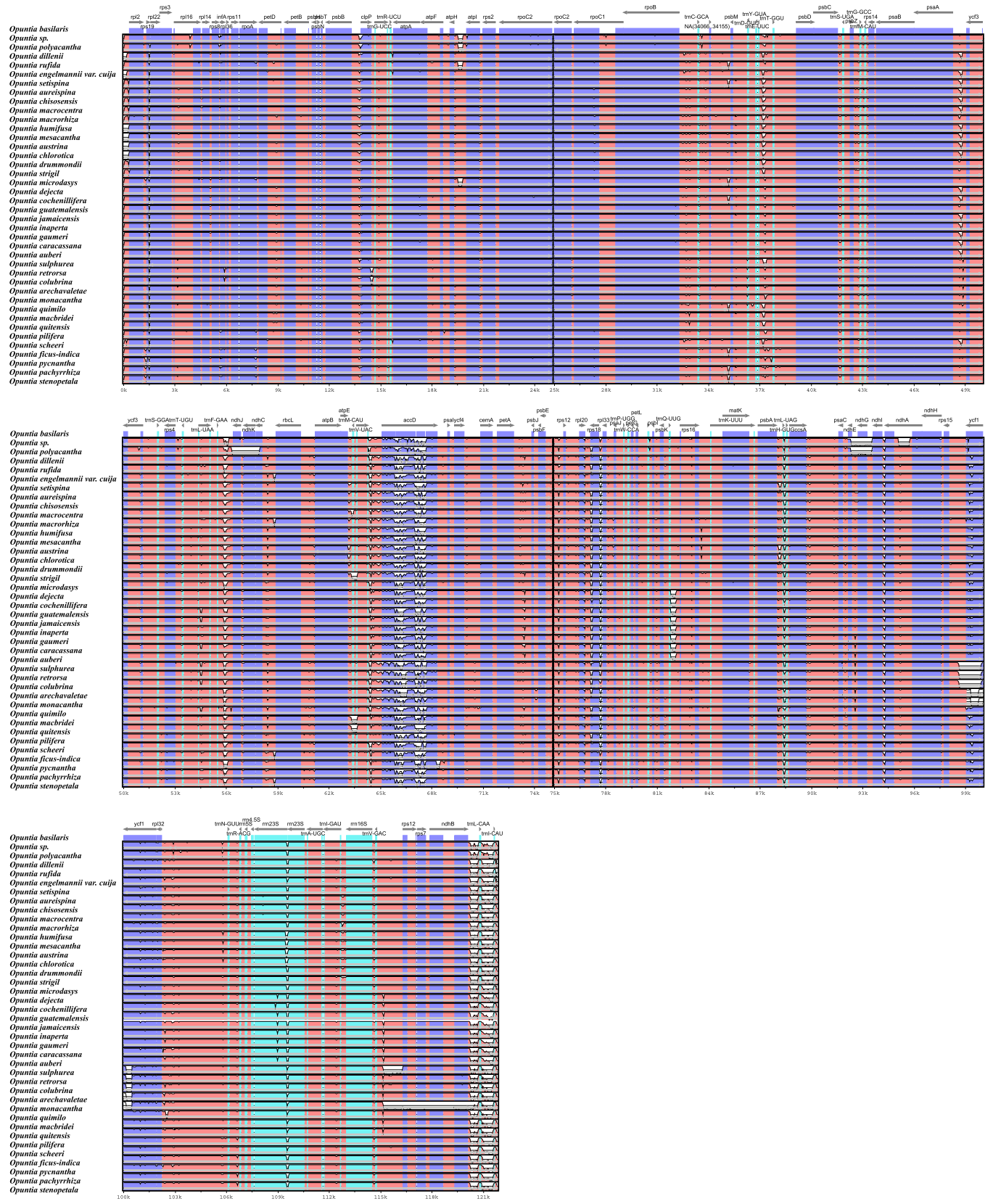


Fig. A3. Comparative visualization of *Opuntia* plastome sequences using mVISTA with *O. basilaris* as a reference. The horizontal lanes display graphs for sequence pairwise identity, where the x-axis represents the base sequence positions and the y-axis indicates the pairwise percent identity, ranging from 50 % to 100 %. Gray arrows denote the orientations and sizes of genes. The color coding within the lane is as follows: dark-blue boxes represent exon regions, light-blue boxes denote tRNA and rRNA regions, and red boxes highlight non-coding sequence regions

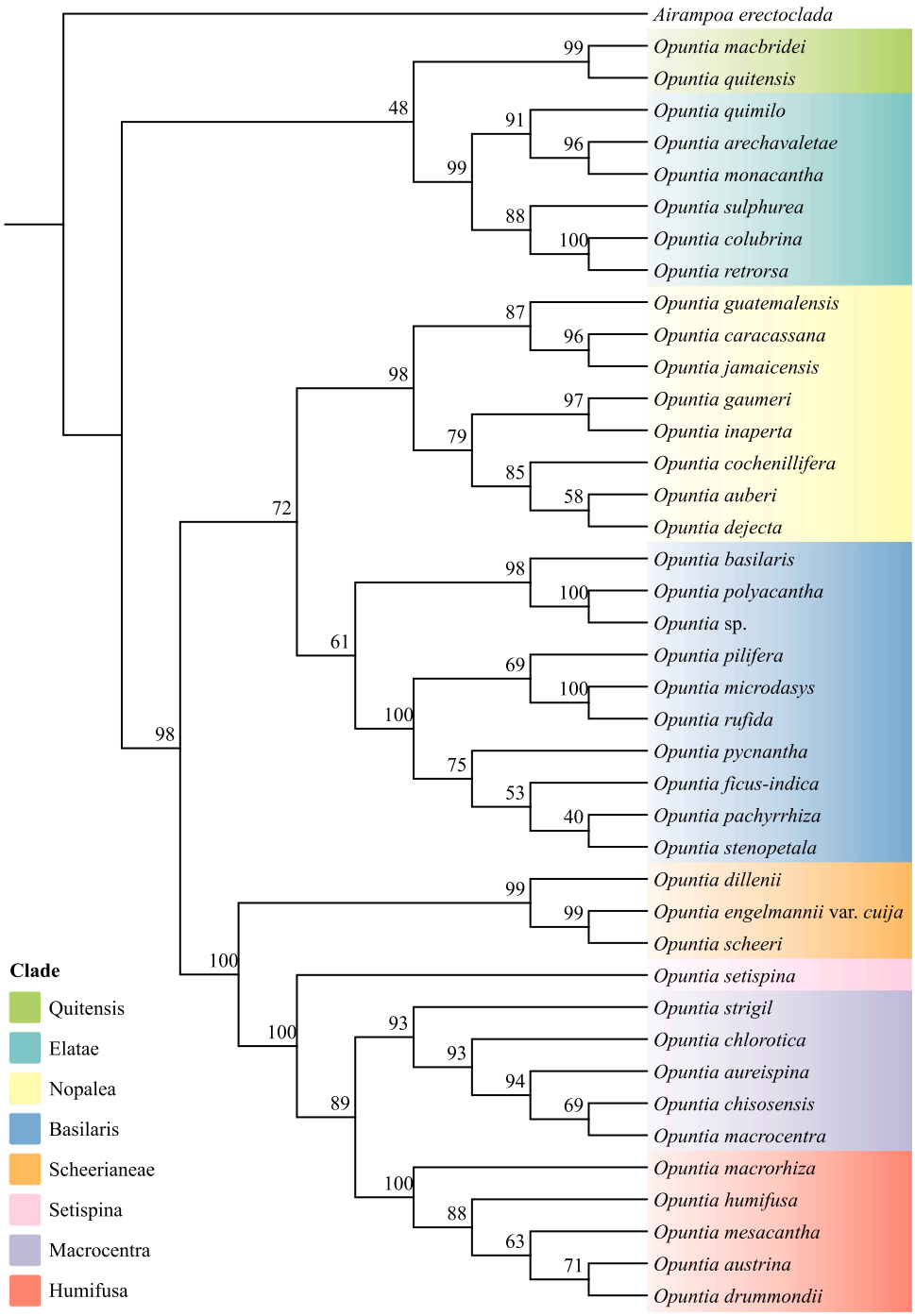


Fig. A4. Maximum Likelihood (ML) phylogenetic tree of *Opuntia* based on six markers (*ndhJ-trnF*, *atpE-trnM-CAU*, *psaJ-rpl33*, *ndhC-rbcL*, *psaA-ycf3*, *atpB-rbcL*). The bootstraps values are shown at each node. The different color blocks represent the eight subclades within the genus *Opuntia*.

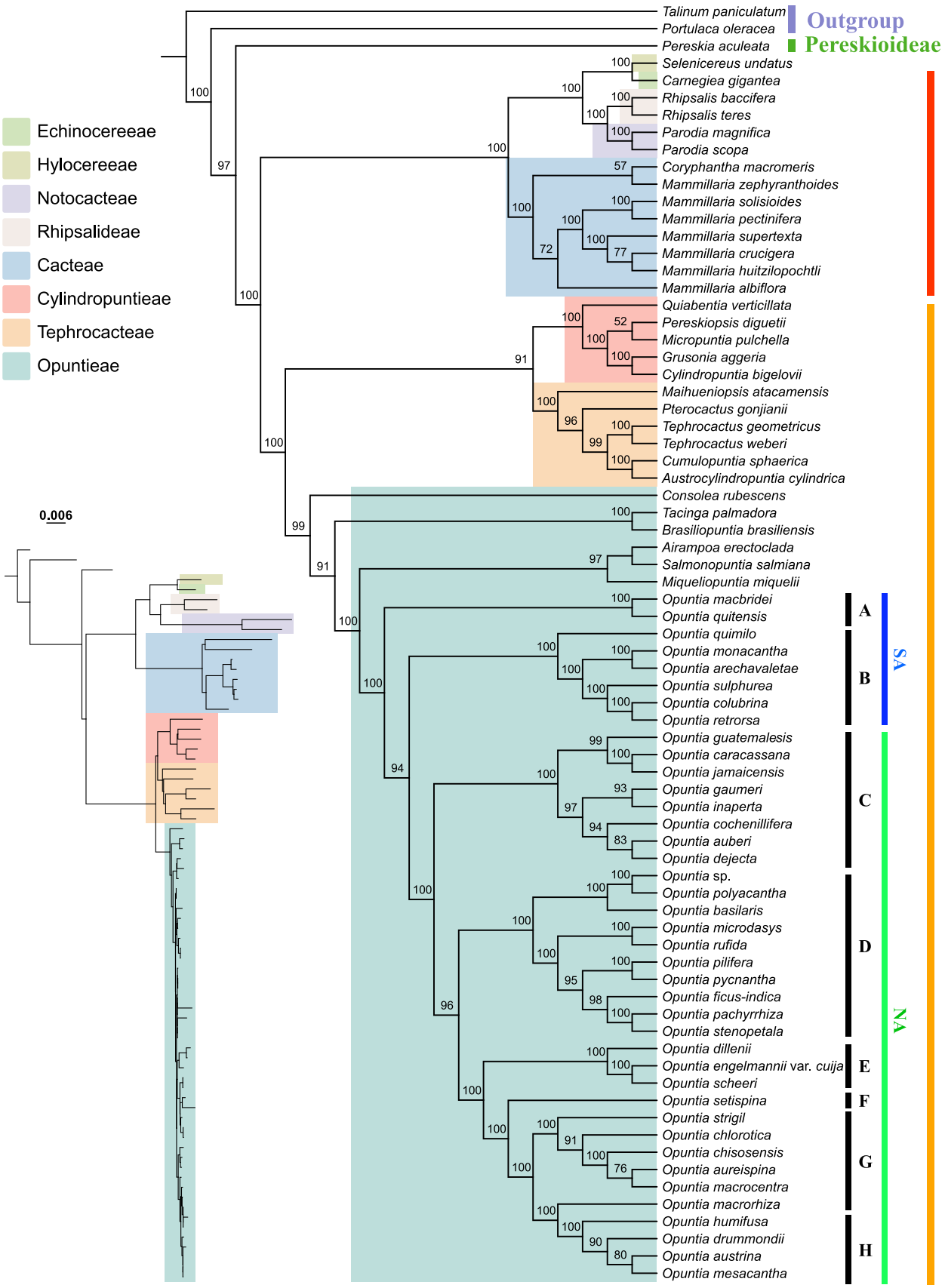


Fig. A5. Maximum likelihood (ML) phylogenetic tree (log-likelihood = –95629.624) inferred using the best-fit model GTR+F+R2, based on 72 plastome protein-coding genes from 72 Cactaceae species and two outgroups. Colored stripes on the right indicate taxonomic groupings: the outgroup and three Cactaceae subfamilies. Within *Opuntia*, blue and green stripes denote the South American (SA) and North American (NA) clades, respectively. Eight major subclades (A–H) are also highlighted. A condensed version of the tree with branch lengths is shown in the lower-left inset

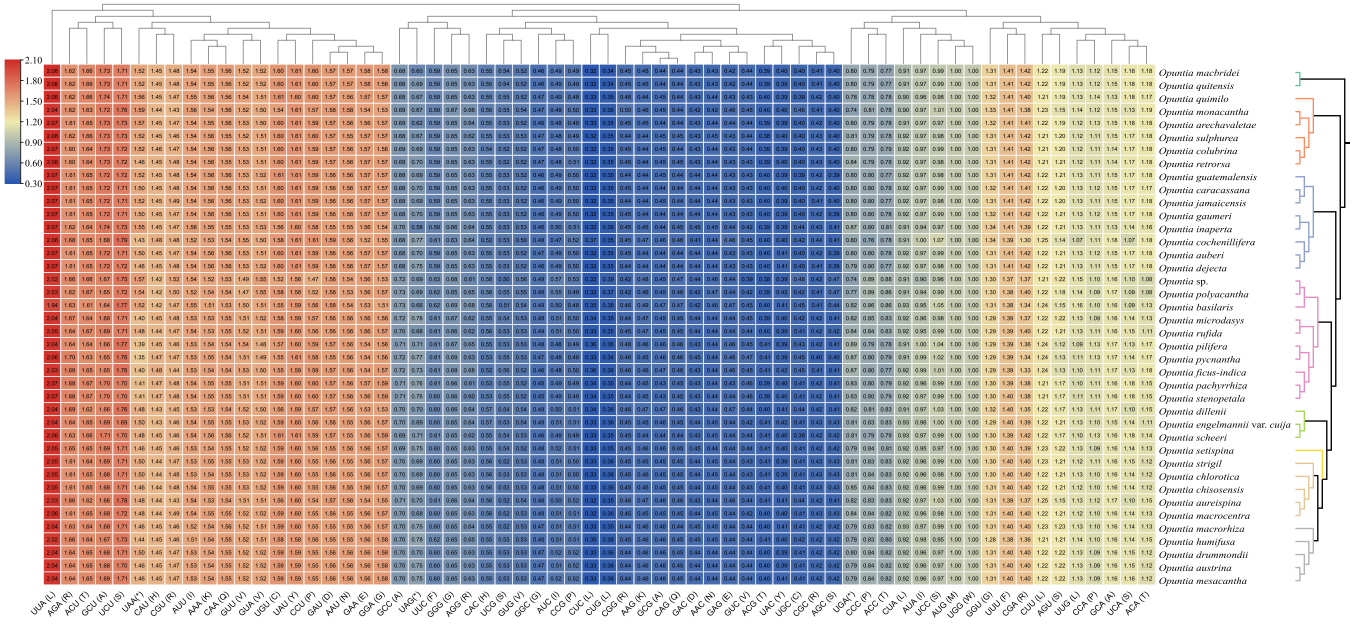


Fig. A6. The heatmap of Relative Synonymous Codon Usage (RSCU) analysis among *Opuntia* plastid CDS

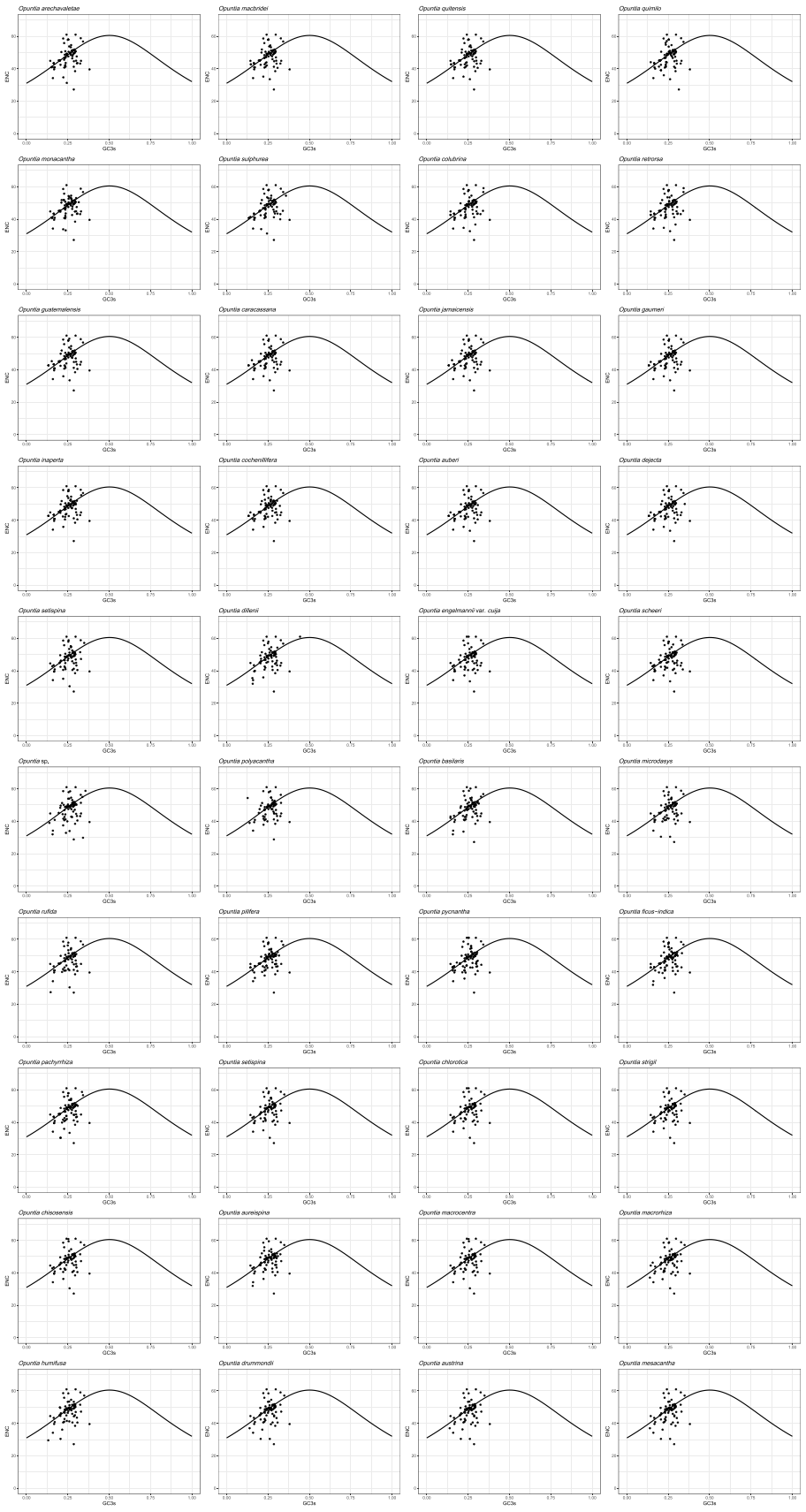


Fig. A7. The Effective Numbers of Codons of CDS among different *Opuntia* plastid

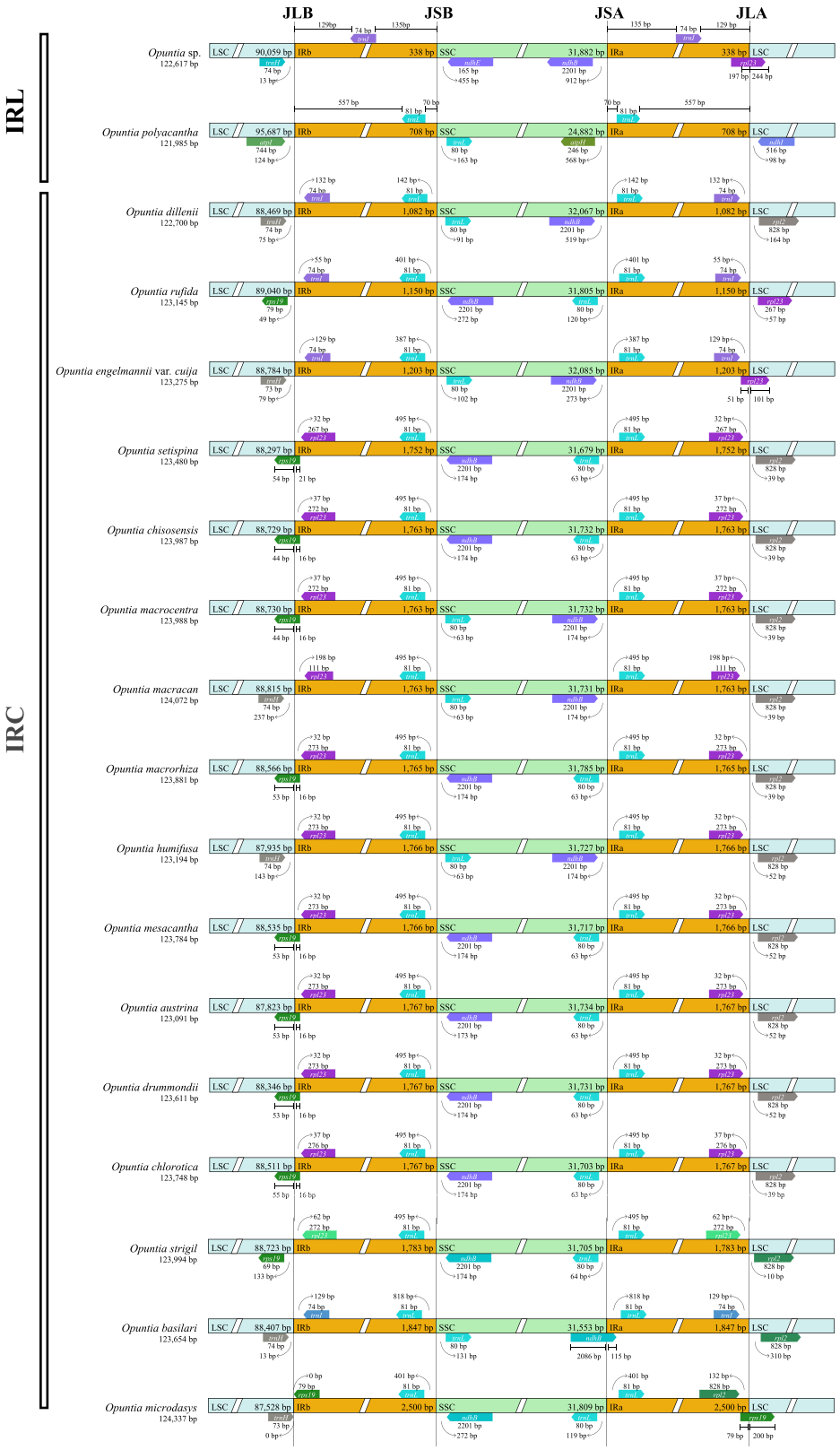


Fig. A8. Comparative analysis of quadripartite structure in *Opuntia* plastomes. The different color blocks represent the four key structural components: the Large Single-Copy (LSC), Inverted Repeat A (IRA), Small Single-Copy (SSC), and Inverted Repeat B (IRB). Arrows indicate the location of genes situated near the boundaries of these structural regions. The junction sites between each region are denoted as JLB (IRB/LSC), JSB (IRB/SSC), JSA (SSC/IRA), and JLA (IRA/LSC), highlighting the specific points of transition between the different segments. The three IR types—IR contracted (IRC), IR normal (IRN), and IR expanded (IRE)—were shown on the left.

Fig. A8. (*continued*).

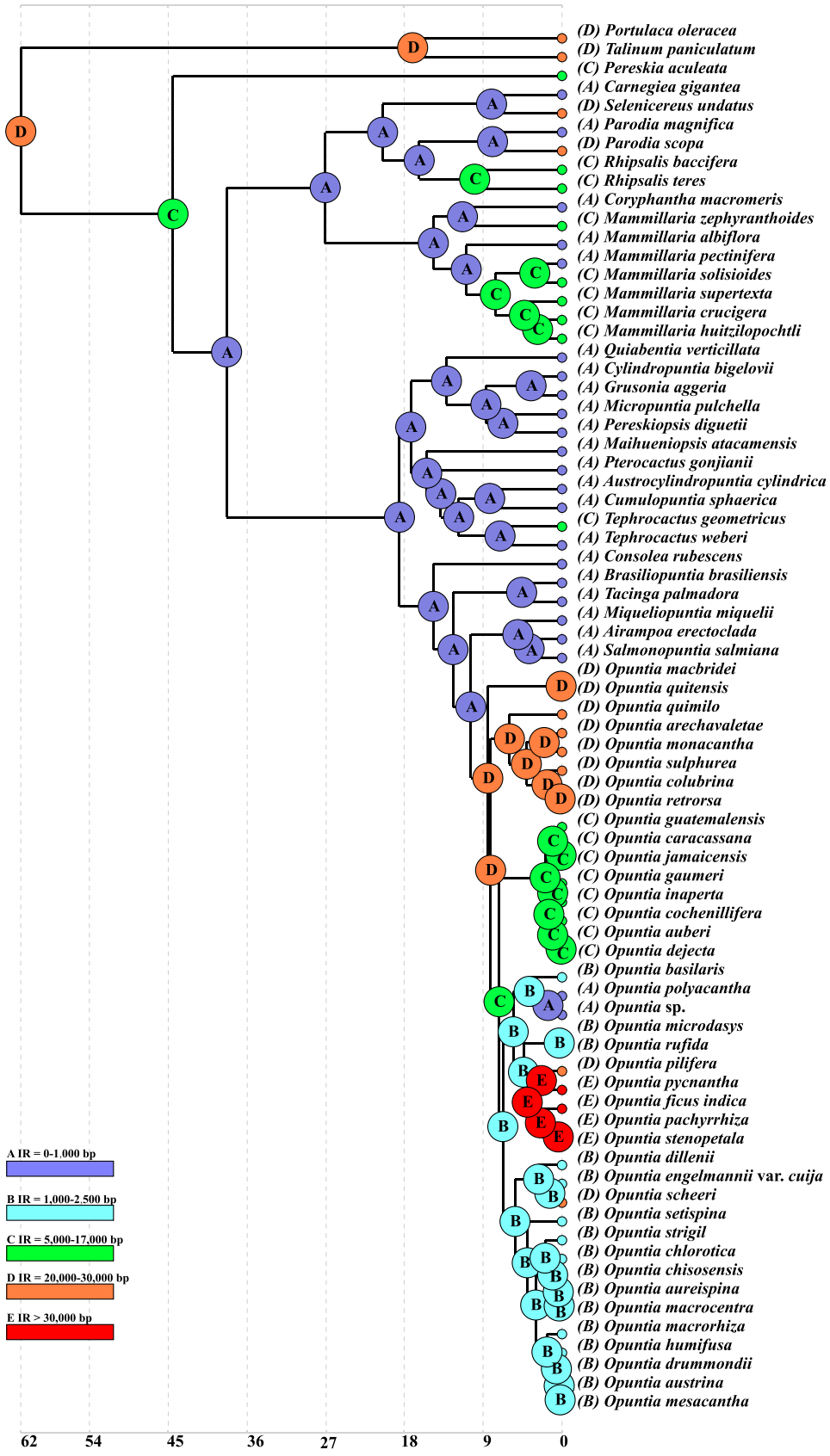


Fig. A9. Ancestral IR state reconstruction in Cactaceae using BayesTraits implemented in RASP. Each node circle indicates the most probable ancestral IR state. Different colors represent varying IR lengths: A) purple for IR 0–1000 bp; B) sky-blue for IR 1000–2500 bp; C) green for IR 25,00–17,000 bp; D) orange for IR 20,000–30,000 bp; E) red for IR more than 30,000 bp. The timeline at the bottom is measured in millions of years ago (Ma)

Appendix B. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.indcrop.2025.121477.

Data availability

Data will be made available on request.

References

Ahmad, N., Nixon, P.J., 2025. Chloroplast alchemy: rewriting the chloroplast genome with high precision. Trends Plant Sci. 30, 341–343. https://doi.org/10.1016/j.tplants.2024.12.014.

Ammar, I., Bardaa, S., Mzid, M., Sahnoun, Z., Rebaii, T., Attia, H., Ennouri, M., 2015. Antioxidant, antibacterial and in vivo dermal wound healing effects of *Opuntia* flower extracts. Int. J. Biol. Macromol. 81, 483–490. https://doi.org/10.1016/j.ijbiomac.2015.08.039.

Anderson, E.F., 2001. The Cactus Family. Timber Press, Portland, USA.

Arakaki, M., Christin, P.-A., Nyffeler, R., Lendel, A., Eggli, U., Ogburn, R.M., Spriggs, E., Moore, M.J., Edwards, E.J., 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. Proc. Natl. Acad. Sci. U. S. A. 108, 8379–8384. https://doi.org/10.1073/pnas.1100628108.

Aruwa, C.E., Amoo, S.O., Kudanga, T., 2018. *Opuntia* (Cactaceae) plant compounds, biological activities and prospects – A comprehensive review. Food Res. Int. 112, 328–344. https://doi.org/10.1016/j.foodres.2018.06.047.

Badawi, G.H., Yamauchi, Y., Shimada, E., Sasaki, R., Kawano, N., Tanaka, Kunisuke, Tanaka, Kiyoshi, 2004. Enhanced tolerance to salt stress and water deficit by overexpressing superoxide dismutase in tobacco (Nicotiana tabacum) chloroplasts. Plant Sci. 166, 919–928. https://doi.org/10.1016/j.plantsci.2003.12.007.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V. M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021.

Barba, F.J., Putnik, P., Bursać Kovačević, D., Poojary, M.M., Roohinejad, S., Lorenzo, J. M., Koubaa, M., 2017. Impact of conventional and non-conventional processing on prickly pear (*Opuntia* spp.) and their derived products: from preservation of beverages to valorization of by-products. Trends Food Sci. Technol. 67, 260–270. https://doi.org/10.1016/j.tifs.2017.07.012.

Beier, S., Thiel, T., Münch, T., Scholz, U., Mascher, M., 2017. Misa-web: a web server for microsatellite prediction. Bioinformatics 33, 2583–2585. https://doi.org/10.1093/bioinformatics/btx198.

Bhattacharya, T., Feng, R., Tierney, J.E., Rubbelke, C., Burls, N., Knapp, S., Fu, M., 2022. Expansion and intensification of the North American monsoon during the pliocene. AGU Adv. 3, e2022AV000757. https://doi.org/10.1029/2022AV000757.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F.K., Müller, N.F., Ogilvie, H.A., Du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., Suchard, M.A., Wu, C.-H., Xie, D., Zhang, C., Stadler, T., Drummond, A.J., 2019. BEAST 2.5: an advanced software platform for bayesian evolutionary analysis. PLoS Comput. Biol. 15, e1006650. https://doi.org/10.1371/journal.pcbi.1006650.

Breslin, P.B., Wojciechowski, M.F., Majure, L.C., 2022. Remarkably rapid, recent diversification of Cochemiea and Mammillaria in the Baja California, Mexico region. Am. J. Bot. 109, 1472–1487. https://doi.org/10.1002/ajb2.16048.

Brudno, M., Malde, S., Poliakov, A., Do, C.B., Couronne, O., Dubchak, I., Batzoglou, S., 2003. Glocal alignment: finding rearrangements during alignment. Bioinformatics 19, i54–i62. https://doi.org/10.1093/bioinformatics/btg1005.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinf. 10, 421. https://doi.org/10.1186/1471-2105-10-421.

Carpena, M., Cassani, L., Gomez-Zavaglia, A., Garcia-Perez, P., Seyyedi-Mansour, S., Cao, H., Simal-Gandara, J., Prieto, M.A., 2023. Application of fermentation for the valorization of residues from Cactaceae family. Food Chem. 410, 135369. https://doi.org/10.1016/j.foodchem.2022.135369.

Cauz-Santos, L.A., da Costa, Z.P., Callot, C., Cauet, S., Zucchi, M.I., Bergès, H., van den Berg, C., Vieira, M.L.C., 2020. A repertory of rearrangements and the loss of an inverted repeat region in *Passiflora* chloroplast genomes. Genome Biol. Evol. 12, 1841–1857. https://doi.org/10.1093/gbe/evaa155.

Chahdoura, H., Barreira, J.C.M., Barros, L., Santos-Buelga, C., Ferreira, I.C.F.R., Achour, L., 2015. Seeds of *Opuntia* spp. as a novel high potential by-product: phytochemical characterization and antioxidant activity. Ind. Crops Prod. 65, 383–389. https://doi.org/10.1016/j.indcrop.2014.11.011.

Chahdoura, H., Barreira, J.C.M., Fernández-Ruiz, V., Morales, P., Calhelha, R.C., Flamini, G., Soković, M., Ferreira, I.C.F.R., Achour, L., 2016. Bioactivity, proximate, mineral and volatile profiles along the flowering stages of *Opuntia microdasys* (Lehm.): defining potential applications. Food Funct. 7, 1458–1467. https://doi.org/10.1039/C5FO01536G.

Challa, S., Neelapu, N.R.R., 2019. Phylogenetic trees: applications, construction, and assessment. In: Hakeem, K.R., Shaik, N.A., Banaganapalli, B., Elango, R. (Eds.), Essentials of Bioinformatics, Volume III: In Silico Life Sciences: Agriculture. Springer International Publishing, Cham, pp. 167–192. https://doi.org/10.1007/978-3-030-19318-8_10.

Chen, J., Zhang, S., Tang, W., Du, X., Yuan, Y., Wu, S., 2022. The Complete Chloroplast Genome Sequence of *Opuntia sulphurea* (Cactaceae). Mitochondrial DNA B 7, 361–362. https://doi.org/10.1080/23802359.2022.2035837.

Choi, I.-S., Jansen, R., Ruhlman, T., 2019. Lost and Found: Return of the Inverted Repeat in the Legume Clade Defined by Its Absence. Genome Biol. Evol. 11, 1321–1333. https://doi.org/10.1093/gbe/evz076.

Chris Blazier, J., Guisinger, M.M., Jansen, R.K., 2011. Recent Loss of Plastid-Encoded *ndh* genes Within *Erodium* (Geraniaceae). Plant Mol. Biol. 76, 263–272. https://doi.org/10.1007/s11103-011-9753-5.

Ciriminna, R., Chavarría-Hernández, N., Rodríguez-Hernández, A.I., Pagliaro, M., 2019. Toward unfolding the bioeconomy of nopal (*Opuntia* spp. Biofuels Bioprod. Biorefin. 13, 1417–1427. https://doi.org/10.1002/bbb.2018.

Comes, H.P., Kadereit, J.W., 1998. The effect of Quaternary climatic changes on plant distribution and evolution. Trends Plant Sci. 3, 432–438. https://doi.org/10.1016/S1360-1385(98)01327-2.

Dalla Costa, T.P., Silva, M.C., De Santana Lopes, A., Gomes Pacheco, T., De Oliveira, J.D., De Baura, V.A., Balsanelli, E., Maltempi De Souza, E., De Oliveira Pedrosa, F., Rogalski, M., 2022. The plastome of *Melocactus glaucescens* Buining & Brederoo reveals unique evolutionary features and loss of essential tRNA genes. Planta 255, 57. https://doi.org/10.1007/s00425-022-03841-2.

Darling, A.E., Mau, B., Perna, N.T., 2010. progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5, e11147. https://doi.org/10.1371/journal.pone.0011147.

Davis, M.B., Shaw, R.G., Etterson, J.R., 2005. Evolutionary responses to changing climate. Ecology 86, 1704–1714. https://doi.org/10.1890/03-0788.

Fages-Lartaud, M., Hundvin, K., Hohmann-Marriott, M.F., 2022. Mechanisms governing codon usage bias and the implications for protein expression in the chloroplast of *Chlamydomonas reinhardtii*. Plant J. 112, 919–945. https://doi.org/10.1111/tpj.15970.

Forsythe, E.S., Williams, A.M., Sloan, D.B., 2021. Genome-wide signatures of plastid-nuclear coevolution point to repeated perturbations of plastid proteostasis systems across angiosperms. Plant Cell 33, 980–997. https://doi.org/10.1093/plcell/koab021.

Galicia-Pérez, A., Golubov, J., Manzanarez-Villasana, G., Martínez-Ramos, L.M., Arias, S., Márquez-Guzmán, J., Mandujano, M.C., 2023. Complex taxonomy in Opuntioideae: is floral morphometry essential to identify *Opuntia* species? Botany 101, 485–497. https://doi.org/10.1139/cjb-2022-0133.

Granados-Aguilar, X., Palomino, G., Martínez-Ramón, J., Arias, S., 2022. Genome evolution and phylogenetic relationships in *Opuntia tehuacana* (Cactaceae, Opuntioideae). Braz. J. Bot. 45, 957–969. https://doi.org/10.1007/s40415-022-00821-4.

Griffith, M.P., 2004. The origins of an important cactus crop, *Opuntia ficus-indica* (Cactaceae): new molecular evidence. Am. J. Bot. 91, 1915–1921. https://doi.org/10.3732/ajb.91.11.1915.

Gu, X., Li, L., Li, S., Shi, W., Zhong, X., Su, Y., Wang, T., 2023. Adaptive evolution and co-evolution of chloroplast genomes in pteridaceae species occupying different habitats: overlapping residues are always highly mutated. BMC Plant Biol. 23, 511. https://doi.org/10.1186/s12870-023-04523-1.

Hernández-Hernández, T., Brown, J.W., Schlumpberger, B.O., Eguiarte, L.E., Magallón, S., 2014. Beyond aridification: multiple explanations for the elevated diversification of cacti in the New World succulent biome. New Phytol. 202, 1382–1397. https://doi.org/10.1111/nph.12752.

Hwang, Y., Han, S., Yoo, C.Y., Hong, L., You, C., Le, B.H., Shi, H., Zhong, S., Hoecker, U., Chen, X., Chen, M., 2022. Anterograde signaling controls plastid transcription via sigma factors separately from nuclear photosynthesis genes. Nat. Commun. 13, 7440. https://doi.org/10.1038/s41467-022-35080-0.

Jin, J., Yu, W., Yang, J., Song, Y., dePamphilis, C.W., Yi, T., Li, D., 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21, 241. https://doi.org/10.1186/s13059-020-02154-5.

John, P., 2004. Correspondence analysis of codon usage [WWW Document]. URL https://codonw.sourceforge.net/culong.html (accessed 4.25.24).

Jorge, A.O.S., Costa, A.S.G., Oliveira, M.B.P.P., 2023. Adapting to climate change with *Opuntia*. Plants 12, 2907. https://doi.org/10.3390/plants12162907.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. Mafft: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066. https://doi.org/10.1093/nar/gkf436.

Kelly, S., 2021. The economics of organellar gene loss and endosymbiotic gene transfer. Genome Biol. 22, 345. https://doi.org/10.1186/s13059-021-02567-w.

Kim, Y.-K., Jo, S., Cheon, S.-H., Joo, M.-J., Hong, J.-R., Kwak, M., Kim, K.-J., 2020. Plastome evolution and phylogeny of orchidaceae, with 24 new sequences. Front. Plant Sci. 11, 22. https://doi.org/10.3389/fpls.2020.00022.

Koch, K.E., Kennedy, R.A., 1980. Effects of seasonal changes in the midwest on crassulacean acid metabolism (CAM) in *Opuntia humifusa* Raf. Oecologia 45, 390–395. https://doi.org/10.1007/BF00540212.

Köhler, M., Reginato, M., Jin, J.-J., Majure, L.C., 2023. More Than a Spiny Morphology: Plastome Variation in the Prickly Pear Cacti (Opuntieae). Ann. Bot. mcad098. https://doi.org/10.1093/aob/mcad098.

Köhler, M., Reginato, M., Souza-Chies, T.T., Majure, L.C., 2020. Insights into chloroplast genome evolution across Opuntioideae (Cactaceae) reveals robust yet sometimes conflicting phylogenetic topologies. Front. Plant Sci. 11 (729). https://doi.org/10.3389/fpls.2020.00729.

Konings, G., 2010. *Opuntia ficus-indica* as an extremophile. Cactus Succul. J. 82, 172–175. https://doi.org/10.2985/015.082.0407.

Korotkova, N., Aquino, D., Arias, S., Eggli, U., Franck, A., Gómez-Hinostrosa, C., Guerrero, P.C., Hernández, H.M., Kohlbecker, A., Köhler, M., Luther, K., Majure, L. C., Müller, A., Metzing, D., Nyffeler, R., Sánchez, D., Schlumpberger, B., Berendsohn, W.G., 2021. Cactaceae at Caryophyllales.org – a dynamic online species-level taxonomic backbone for the family. Willdenowia 51, 251–270. https://doi.org/10.3372/wi.51.51208.

Krämer, C., Boehm, C.R., Liu, J., Ting, M.K.Y., Hertle, A.P., Forner, J., Ruf, S., Schöttler, M.A., Zoschke, R., Bock, R., 2024. Removal of the large inverted repeat from the plastid genome reveals gene dosage effects and leads to increased genome copy number. Nat. Plants 10, 923–935. https://doi.org/10.1038/s41477-024-01709-9.

Kumar, D., Sharma, P.K., 2020. A review on *Opuntia* species and its chemistry, pharmacognosy, pharmacology and bioapplications. Curr. Nutr. Food Sci. 16, 1227–1244. https://doi.org/10.2174/1573401316666200220092414.

Labra, M., Grassi, F., Bardini, M., Imazio, S., Guiggi, A., Citterio, S., Banfi, E., Sgorbati, S., 2003. Genetic relationships in *Opuntia* Mill. genus (Cactaceae) detected by molecular marker. Plant Sci. 165, 1129–1136. https://doi.org/10.1016/S0168-9452(03)00321-2.

Labs, M., Rühle, T., Leister, D., 2016. The antimycin A-snsitive pathway of cyclic electron flow: from 1963 to 2015. Photosynth. Res. 129, 231–238. https://doi.org/10.1007/s11120-016-0217-2.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Las Casas, G., Distefano, G., Caruso, M., Nicolosi, E., Gentile, A., La Malfa, S., 2018. Relationships among cultivated *Opuntia ficus-indica* genotypes and related species assessed by cytoplasmic markers. Genet. Resour. Crop Evol. 65, 759–773. https://doi.org/10.1007/s10722-017-0569-2.

Lee, C., Choi, I.-S., Cardoso, D., de Lima, H.C., de Queiroz, L.P., Wojciechowski, M.F., Jansen, R.K., Ruhlman, T.A., 2021. The chicken or the egg? plastome evolution and an independent loss of the inverted repeat in papilionoid legumes. Plant J. 107, 861–875. https://doi.org/10.1111/tpj.15351.

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. https://doi.org/10.48550/arXiv.1303.3997.

Li, S., Chang, L., Zhang, J., 2021. Advancing organelle genome transformation and editing for crop improvement. Plant Commun. 2. https://doi.org/10.1016/j.xplc.2021.100141.

Li, H., Guo, Q., Xu, L., Gao, H., Liu, L., Zhou, X., 2023. Cpjsdraw: analysis and visualization of junction sites of chloroplast genomes. PeerJ 11, e15326. https://doi.org/10.7717/peerj.15326.

Lin, C.S., Chen, J.J.W., Chiu, C., Hsiao, H.C.W., Yang, C., Jin, X., Leebens-Mack, J., de Pamphilis, C.W., Huang, Y., Yang, L., Chang, W., Kui, L., Wong, G.K.S., Hu, J., Wang, W., Shih, M., 2017. Concomitant loss of NDH complex-related genes within chloroplast and nuclear genomes in some orchids. Plant J. 90, 994–1006. https://doi.org/10.1111/tpj.13525.

Lin, C.-S., Chen, J.J.W., Huang, Y., Chan, M.T., Daniell, H., Chang, W., Hsu, C.T., Liao, D. C., Wu, F.H., Lin, S.Y., Liao, C., Deyholos, M.K., Wong, G.K.S., Albert, V.A., Chou, M., Chen, C., Shih, M., 2015. The location and translocation of *ndh* genes of chloroplast origin in the Orchidaceae family. Sci. Rep. 5, 9040. https://doi.org/10.1038/srep09040.

Liu, S., Xue, D., Cheng, R., Han, H., 2014. The complete mitogenome of *Apocheima cinerarius* (Lepidoptera: Geometridae: Ennominae) and comparison with that of other lepidopteran insects. Gene 547, 136–144. https://doi.org/10.1016/j.gene.2014.06.044.

Lu, W.C., Chiu, C.S., Chan, Y., Mulio, A.T., Li, P.-H., 2023. Recent research on different parts and extracts of *Opuntia dillenii* and its bioactive components, functional properties, and applications. Nutrients 15, 2962. https://doi.org/10.3390/nu15132962.

Majure, L.C., Achá, S., Baker, M.A., Puente-Martínez, R., Köhler, M., Fehlberg, S., 2023. Phylogenomics of one of the world's most intriguing groups of CAM plants, the opuntioids (Opuntioideae: Cactaceae): adaptation to tropical dry forests helped drive prominent morphological features in the clade. Diversity 15, 570. https://doi.org/10.3390/d15040570.

Majure, L.C., Judd, W.S., Soltis, P.S., Soltis, D.E., 2017. Taxonomic Revision of the *Opuntia humifusa* complex (Opuntieae: Cactaceae) of the Eastern United States. Phytotaxa 290 (1). https://doi.org/10.11646/phytotaxa.290.1.1.

Majure, L.C., Puente, R., Griffith, M.P., Judd, W.S., Soltis, P.S., Soltis, D.E., 2012. Phylogeny of *Opuntia* s.s. (Cactaceae): clade delineation, geographic origins, and reticulate evolution. Am. J. Bot. 99, 847–864. https://doi.org/10.3732/ajb.1100375.

Mao, K.S., Wang, Y., Liu, J.Q., 2021. Evolutionary origin of species diversity on the Qinghai–Tibet Plateau. J. Syst. Evol. 59, 1142–1158. https://doi.org/10.1111/jse.12809.

Maréchal, A., Brisson, N., 2010. Recombination and the maintenance of plant organelle genome stability. New Phytol. 186, 299–317. https://doi.org/10.1111/j.1469-8137.2010.03195.x.

Maréchal, A., Parent, J.-S., Véronneau-Lafortune, F., Joyeux, A., Lang, B.F., Brisson, N., 2009. Whirly proteins maintain plastid genome stability in arabidopsis. Proc. Natl. Acad. Sci. USA 106, 14693–14698. https://doi.org/10.1073/pnas.0901710106.

Mason, P.M., Glover, K., Smith, J.A.C., Willis, K.J., Woods, J., Thompson, I.P., 2015. The potential of CAM crops as a globally significant bioenergy resource: moving from 'Fuel or Food' to 'Fuel and more Food. Energy Environ. Sci. 8, 2320–2329. https://doi.org/10.1039/C5EE00242G.

Meade, A., Pagel, M., 2018. Bayestraits: a computer package for analyses of trait evolution. ⟨https://www.evolution.reading.ac.uk/BayesTraitsV4.1.1/BayesTraitsV4.1.1.html⟩.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

Mohanta, T.K., Mishra, A.K., Khan, A., Hashem, A., Abd_Allah, E.F., Al-Harrasi, A., 2020. Gene loss and evolution of the plastome. Genes 11, 1133. https://doi.org/10.3390/genes11101133.

Nyffeler, R., 2002. Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from *trnK/ matK* and *trnL-trnF* sequences. Am. J. Bot. 89, 312–326. https://doi.org/10.3732/ajb.89.2.312.

Pinkava, D.J., 2003. Cactaceae Cactus Family, Part Six. *Opuntia* P. Miller Prickly-Pears. J. Ariz. Nev. Acad. Sci. 35, 137–150.

Piot, A., Hackel, J., Christin, P.-A., Besnard, G., 2018. One-third of the plastid genes evolved under positive selection in PACMAD grasses. Planta 247, 255–266. https://doi.org/10.1007/s00425-017-2781-x.

Qi, Y., Xu, W., Xing, T., Zhao, M., Li, N., Yan, L., Xia, G., Wang, M., 2015. Synonymous codon usage bias in the plastid genome is unrelated to gene structure and shows evolutionary heterogeneity. Evol. Bioinf. 11, EBO.S22566. https://doi.org/10.4137/EBO.S22566.

Qu, X., Moore, M., Li, D., Yi, T., 2019. PGA: A software package for rapid, accurate, and flexible batch annotation of plastomes. Plant Methods 15, 1–12. https://doi.org/10.1186/s13007-019-0435-7.

Quintanar-Orozco, E.T., Vázquez-Rodríguez, G.A., Beltrán-Hernández, R.I., Lucho-Constantino, C.A., Coronel-Olivares, C., Montiel, S.G., Islas-Valdez, S., 2018. Enhancement of the biogas and biofertilizer production from *Opuntia heliabravoana* Scheinvar. Environ. Sci. Pollut. Res. 25, 28403–28412. https://doi.org/10.1007/s11356-018-2845-x.

Rabah, S.O., Shrestha, B., Hajrah, N.H., Sabir, Mumdooh J., Alharby, H.F., Sabir, Mernan J., Alhebshi, A.M., Sabir, J.S.M., Gilbert, L.E., Ruhlman, T.A., Jansen, R.K., 2019. *Passiflora* plastome sequencing reveals widespread genomic rearrangements. J. Syst. Evol. 57, 1–14. https://doi.org/10.1111/jse.12425.

Rache-Cardenal, L.Y., Albesiano-Hoyos, A.S., Tall, H., 2022. Optimization of a DNA extraction method using subepidermis from *Austrocylindropuntia* and *Opuntia*-Opuntioideae. Rev. Investig. Desarro. Innov. 12, 305–314.

Ramadan, M.F., Ayoub, T.E.M., Rohn, S., 2021. *Opuntia* spp.: Chemistry, Bioactivity and Industrial Applications. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-78444-7.

Revell, L.J., 2012. Phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3, 217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x.

Rodrigues, C., Paula, C.D., de, Lahbouki, S., Meddich, A., Outzourhit, A., Rashad, M., Pari, L., Coelhoso, I., Fernando, A.L., Souza, V.G.L., 2023. *Opuntia* spp.: an overview of the bioactive profile and food applications of this versatile crop adapted to arid lands. Foods 12, 1465. https://doi.org/10.3390/foods12071465.

Rousseau-Gueutin, M., Huang, X., Higginson, E., Ayliffe, M., Day, A., Timmis, J.N., 2013. Potential functional replacement of the plastidic acetyl-CoA carboxylase subunit (*accD*) gene by recent transfers to the nucleus in some angiosperm lineages. Plant Physiol. 161, 1918–1929. https://doi.org/10.1104/pp.113.214528.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sánchez-Gracia, A., 2017. Dnasp 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34, 3299–3302. https://doi.org/10.1093/molbev/msx248.

Sanderson, M.J., Copetti, D., Burquez, A., Bustamante, E., Charboneau, J.L.M., Eguiarte, L.E., Kumar, S., Lee, H.O., Lee, J., McMahon, M., Steele, K., Wing, R., Yang, T.J., Zwickl, D., Wojciechowski, M.F., 2015. Exceptional reduction of the plastid genome of saguaro cactus (*Carnegiea gigantea*): loss of the *ndh* gene suite and inverted repeat. Am. J. Bot. 102, 1115–1127. https://doi.org/10.3732/ajb.1500184.

Santos-Díaz, M.S., Barba de la Rosa, A.-P., Héliès-Toussaint, C., Guéraud, F., Nègre-Salvayre, A., 2017. *Opuntia* spp.: characterization and benefits in chronic diseases. Oxid. Med. Cell. Longev. 2017, e8634249. https://doi.org/10.1155/2017/8634249.

Shoukat, R., Cappai, M., Pia, G., Pilia, L., 2023. An updated review: *Opuntia ficus indica* (OFI) chemistry and its diverse applications. Appl. Sci. 13, 7724. https://doi.org/10.3390/app13137724.

Sniderman, J.M.K., Woodhead, J.D., Hellstrom, J., Jordan, G.J., Drysdale, R.N., Tyler, J. J., Porch, N., 2016. Pliocene reversal of late Neogene aridification. Proc. Natl. Acad. Sci. USA 113, 1999–2004. https://doi.org/10.1073/pnas.1520188113.

Solórzano, S., Chincoya, D.A., Sanchez-Flores, A., Estrada, K., Díaz-Velásquez, C.E., González-Rodríguez, A., Vaca-Paniagua, F., Dávila, P., Arias, S., 2019. De novo assembly discovered novel structures in genome of plastids and revealed divergent inverted repeats in *Mammillaria* (Cactaceae, Caryophyllales). Plants 8, 392. https://doi.org/10.3390/plants8100392.

Strand, D.D., D'Andrea, L., Bock, R., 2019. The plastid NAD (P) H dehydrogenase-like complex: structure, function and evolutionary dynamics. Biochem. J. 476, 2743–2756.

Timmis, J.N., Ayliffe, M.A., Huang, C.Y., Martin, W., 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat. Rev. Genet. 5, 123–135. https://doi.org/10.1038/nrg1271.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S. G., 2012. Primer3—new capabilities and interfaces. Nucleic Acids Res. 40. https://doi.org/10.1093/nar/gks596 e115–e115.

Valadez-Moctezuma, E., Samah, S., Mascorro-Gallardo, J.O., Marbán-Mendoza, N., Aranda-Osorio, G., Flores-Girón, E., Brito-Nájera, G., Rodríguez de la O, J.L., 2022. The first transcriptomic analyses of fruits and cladodes for comparison between three species of *Opuntia*. Genet. Resour. Crop Evol. https://doi.org/10.1007/s10722-022-01480-w.

Velasco-de León, M.P., Spicer, R.A., Steart, D.C., 2010. Climatic reconstruction of two Pliocene floras from Mexico. Palaeobiodivers. Palaeoenviron 90, 99–110. https://doi.org/10.1007/s12549-010-0022-4.

Vergara-Cruces, Á., Pramanick, I., Pearce, D., Vogirala, V.K., Byrne, M.J., Low, J.K.K., Webster, M.W., 2024. Structure of the plant plastid-encoded RNA polymerase. Cell 187, 1145–1159.e21. https://doi.org/10.1016/j.cell.2024.01.036.

Wang, Z.-X., Wang, D.-J., Yi, T.-S., 2022. Does IR-loss promote plastome structural variation and sequence evolution? Front. Plant Sci. 13, 888049. https://doi.org/10.3389/fpls.2022.888049.

Wang, T., Wang, G.-L., Fang, Y., Zhang, Y., Peng, W., Zhou, Y., Zhang, A., Yu, L.-J., Lu, C., 2024. Architecture of the spinach plastid-encoded RNA polymerase. Nat. Commun. 15, 9838. https://doi.org/10.1038/s41467-024-54266-2.

Wang, D., Zhang, Y., Zhang, Z., Zhu, J., Yu, J., 2010. Kaks_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genom. Proteom. Bioinforma. 8, 77–80. https://doi.org/10.1016/S1672-0229(10)60008-3.

Wei, N., Pérez-Escobar, O.A., Musili, P.M., Huang, W.-C., Yang, J.-B., Hu, A.-Q., Hu, G.-W., Grace, O.M., Wang, Q.-F., 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, 712064. https://doi.org/10.3389/fpls.2021.712064.

Weng, M.-L., Ruhlman, T.A., Jansen, R.K., 2016. Plastid–nuclear interaction and accelerated coevolution in plastid ribosomal genes in geraniaceae. Genome Biol. Evol. 8, 1824–1838. https://doi.org/10.1093/gbe/evw115.

Wick, R.R., Schultz, M.B., Zobel, J., Holt, K.E., 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 31, 3350–3352. https://doi.org/10.1093/bioinformatics/btv383.

Wicke, S., Schneeweiss, G.M., dePamphilis, C.W., Müller, K.F., Quandt, D., 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol. Biol. 76, 273–297. https://doi.org/10.1007/s11103-011-9762-4.

Wickham, H., 2011. Ggplot2. WIREs Comput. Stat 3, 180–185. https://doi.org/10.1002/wics.147.

Wu, S., Chen, J., Li, Y., Liu, A., Li, A., Yin, M., Shrestha, N., Liu, J., Ren, G., 2021. Extensive Genomic Rearrangements Mediated by Repetitive Sequences in Plastomes of *Medicago* and Its Relatives. BMC Plant Biol. 21, 421. https://doi.org/10.1186/s12870-021-03202-3.

Wu, H., Li, D.-Z., Ma, P.-F., 2024. Unprecedented variation pattern of plastid genomes and the potential role in adaptive evolution in poales. BMC Biol. 22, 97. https://doi.org/10.1186/s12915-024-01890-5.

Xie, J., Chen, Y., Cai, G., Cai, R., Hu, Z., Wang, H., 2023. Tree visualization by one table (tvBOT): a web application for visualizing, modifying and annotating phylogenetic trees. Nucleic Acids Res gkad359. https://doi.org/10.1093/nar/gkad359.

Yahia, E., Saenz, C., 2011. 14 - Cactus pear (*Opuntia* species). In: Yahia, E. (Ed.), Postharvest Biology and Technology of Tropical and Subtropical Fruits. Woodhead Publishing Series in Food Science, Technology and Nutrition. Woodhead Publishing, pp. 290–329. https://doi.org/10.1533/9780857092762.290.

Yamamoto, H., Peng, L., Fukao, Y., Shikanai, T., 2011. An src homology 3 domain-like fold protein forms a ferredoxin binding site for the chloroplast NADH dehydrogenase-like complex in *Arabidopsis*. Plant Cell 23, 1480–1493. https://doi.org/10.1105/tpc.110.080291.

Yang, X., Liu, D., Tschaplinski, T.J., Tuskan, G.A., 2019. Comparative genomics can provide new insights into the evolutionary mechanisms and gene function in CAM plants. J. Exp. Bot. 70, 6539–6547. https://doi.org/10.1093/jxb/erz408.

Yao, G., Jin, J., Li, H., Yang, J., Mandala, V.S., Croley, M., Mostow, R., Douglas, N.A., Chase, M.W., Christenhusz, M.J.M., Soltis, D.E., Soltis, P.S., Smith, S.A., Brockington, S.F., Moore, M.J., Yi, T., Li, D., 2019. Plastid phylogenomic insights into the evolution of Caryophyllales. Mol. Phylogenet. Evol. 134, 74–86. https://doi.org/10.1016/j.ympev.2018.12.023.

Yu, Y., Harris, A.J., Blair, C., He, X., 2015. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. Mol. Phylogenet. Evol. 87, 46–49. https://doi.org/10.1016/j.ympev.2015.03.008.

Yu, J., Li, J., Zuo, Y., Qin, Q., Zeng, S., Rennenberg, H., Deng, H., 2023. Plastome variations reveal the distinct evolutionary scenarios of plastomes in the subfamily Cereoideae (Cactaceae). BMC Plant Biol. 23, 132. https://doi.org/10.1186/s12870-023-04148-4.

Yu, J., Ran, Z., Zhang, J., Wei, L., Ma, W., 2022. Genome-wide insights into the organelle translocation of photosynthetic NDH-1 genes during evolution. Front. Microbiol. 13, 956578. https://doi.org/10.3389/fmicb.2022.956578.

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.-Y., 2017. GGTREE: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 8, 28–36. https://doi.org/10.1111/2041-210X.12628.

Zeghbib, W., Boudjouan, F., Vasconcelos, V., Lopes, G., 2022. Phenolic compounds' occurrence in *Opuntia* species and their role in the inflammatory process: a review. Molecules 27, 4763. https://doi.org/10.3390/molecules27154763.

Zhang, Y., Tian, L., Lu, C., 2023. Chloroplast gene expression: recent advances and perspectives. Plant Commun. 4. https://doi.org/10.1016/j.xplc.2023.100611.

Zhang, P., Xu, W., Lu, X., Wang, L., 2021. Analysis of codon usage bias of chloroplast genomes in *Gynostemma* species. Physiol. Mol. Biol. Plants 27, 2727–2737. https://doi.org/10.1007/s12298-021-01105-z.

Zhu, A., Guo, W., Gupta, S., Fan, W., Mower, J.P., 2016. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. N. Phytol. 209, 1747–1756. https://doi.org/10.1111/nph.13743.