Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Insights into the multi-chromosomal mitochondrial genome structure of the Xero- halophytic plant Haloxylon ammodendron (C.A.Mey.) Bunge ex Fenzl

Lulu Yang
Shenzhen University Medical School, Shenzhen University
Wenjun Guo
Sun Yat-sen University Cancer Center
Zehan Zheng
Shenzhen University Medical School, Shenzhen University
Tian Xiao (txiao@szu.edu.cn)
Shenzhen University Medical School, Shenzhen University

Research Article

Keywords: Haloxylon, xerophytic plant, mitochondrial genome, multi-chromosome, DNA migration

Posted Date: September 20th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3305263/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Genomics on January 29th, 2024. See the published version at https://doi.org/10.1186/s12864-024-10026-6.

Abstract

Background

Haloxylon ammodendron is an important ecological plant due to its adaptation ability in desert conditions, halophytic environments, and the fixation of sand. It is a promising plant with the ecological value that could help carbon sequestration and a valuable C4 plant model for elucidating the genetic basis of extreme drought tolerance. Although plastid and nuclear genomes were made available, the lack of a mitochondrial genome hampered comprehensive insights into its mitogenome structure organization and phylogenetic implications.

Results

In the present study, the mitochondrial genome of *H. ammodendron* was assembled and annotated, resulting in a multi-chromosomal configuration with two circular chromosomes. The mitogenome was 210,149 bp in length and contained 31 protein-coding genes, 18 tRNA, and three rRNA. A total of 66 simple sequence repeats (SSRs) were identified. Besides, 27 tandem repeats, 312 forward repeats, and 303 palindromic repeats were found. A set of 17 sequence fragments were found homologous between the mitochondrial and chloroplast genomes, spanning 5,233 bp, accounting for 2.49% of the total mitogenome size. In total, 337 RNA editing sites were predicted, all C-to-U conversion types. Phylogenetic inference confidently placed *H. ammodendron* in the Amaranthacea family and its close relative, *Suaeda glacum*.

Conclusions

H. ammodendron mitogenome showed a multi-chromosome structure with two fully circularized molecules. The present mitogenome represents a valuable resource for understanding the basis of mitogenome structure variation within Caryophyllales and evolution in land plants, their identification, and classification.

Background

H. ammodendron is a small C4 perennial tree belonging to the Amaranthaceae family, formerly Chenopodiaceae [1]. The genus *Haloxylon* encompasses approximately 11 species, with a distribution area ranging from Iran to Mongolia and Northwestern China [2]. *H. ammodendron* termed the King of psammophytic woody plants, has remarkable adaptability to severe environmental conditions, including desert conditions and sandstorms [3]. It is widely found in the deserts of northwest China and plays a crucial role in environmental protection through wind control, carbon sequestration, and sand fixation [4]. However, due to increasing overexploitation and the dramatic climate change occurring in central Asia, the *Haloxylon* plants have been threatened [5].

These plants' declining numbers and environmental significance have sparked significant interest in genetic research on *Haloxylon* resources. Long et al. [6] provided the first transcriptome resource of *H. ammodendron* to investigate genes that confer drought stress resistance to the plant. For a thorough understanding of *H. ammodendron*'s resilience to drought stress, Gao et al. [7]conducted a complete transcriptome analysis under osmotic conditions, paired with evaluations of physiological factors like proline, betaine, soluble sugar levels, and peroxidase activity. Using chloroplast genes (*trnS-trnG* and *trnV*) and ITS (Internal transcribed spacer) (ITS1-ITS4) sequences, Chen et al. [8] assessed 420 individuals from 36 populations of *Haloxylon* for their genetic diversity and historical demography. Suo et al. [9] utilized SSR marker-derived DNA markers to distinguish *H. ammodendron* and *Haloxylon persicum* species. Later, the chloroplast genomes of *H. ammodendron* and *H. persicum* were assembled using short-reads sequencing data. Recently, a high-quality chromosome-level of *H. ammodendron* was released [10], paving the way for an in-depth investigation of genes of interest concerning its high tolerance to drought and salinity. Despite the availability of these genomic resources, the mitochondrial genome of *H. ammodendron* is lacking, hampering mitogenome-based structural and functional evolutionary studies.

The mitogenome, often called the cell's powerhouse, plays a crucial role in cellular energy production [11]. Housed in the mitochondria, this distinct genome is responsible for producing crucial proteins for oxidative phosphorylation, which leads to the creation of adenosine triphosphate (ATP) – the primary energy currency of the cell [12, 13]. High variations in mitogenome size coupled with structural complexity are the major factors slowing the pace of fully and accurately assembled mitogenome projects. For example, mitogenome sizes vary over 40-fold in the single genus Silene, from 253 kbp in *Silene latifolia* to more than 11 Mbp in *Silene conica* [14]. Variation of repetitive sequences, incorporation of foreign sequences, and gain or loss of non-coding regions are among the key factors contributing to genome size variability [15, 16].

Furthermore, while a single circular mitogenome is common in angiosperms, various configurations may occur, including linear, fragmented, loop-like, branched, and multi-chromosomal [11]. The dynamic nature of mitochondrial genomes of present challenges in fully assembling the genomes of certain species compared to the chloroplast genome assembly.

However, advancements in long-read sequencing technologies have made the assembly of plant mitochondrial genomes more achievable, facilitating the comparative studies of plant mitogenome structure and evolutionary implications [17, 18]. In the present study, short and long reads sequencing data were used to generate the first mitochondrial genome of *H. ammodendron*. The characterization of repeats content, sequence collinearity, DNA migration, and RNA editing sites were explored.

Results

Genome assembly and annotation

In this study, we utilized Illumina short-reads and PacBio HiFi long-reads to assemble the mitochondrial genome of *H. ammodendron*. Importantly, the data was not generated within this study; instead, we sourced the Illumina short-reads and PacBio long-reads from a previously conducted study on the nuclear genome of the species. Specifically, we accessed 49.4 Gbp of clean short reads (Table S1) from the NCBI SRA database under accession number SRR17127859. The assembly was facilitated using the GetOrganelle software, version 1.7.5 [19]. In parallel, the PacBio dataset comprised 22.6 Gbp of long-reads, also retrieved from NCBI SRA under the accession SRR17129371. When visualizing the short-read assembly graph using Bandage version 0.8.1 [20], an intricate and circular multi-branch mitochondrial genome structure was evident, encompassing nine distinct nodes or contigs (as illustrated in Fig. 1A). These nodes were delineated by overlapping regions depicted by black lines on the graph. Closer manual scrutiny allowed us to discern the two pseudo-chromosomes of the mitogenome (Fig. 1B). The sequence for path resolution ordering is elaborated in Table S2. Overlapping regions were addressed by aligning associated branch nodes with the PacBio HiFi dataset. The culmination of this effort was the acquisition of two fully circular sequences that represent the mitochondrial genome of *H. ammodendron*, as displayed in Fig. 1C. The two circularized contigs, namely Chromosome 1 and Chromosome 2, exhibited similar GC content (44%) spanning 121,403 bp and 88,746 bp, respectively (Table 1).

Table 1
Basic statistics of *H. ammodendron* mitochondrial genome

Contigs	Type	Length	GC content
Chromosome 1–2	Branched	210,149 bp	44.42%
Chromosome 1	circular	121,403 bp	44.58%
Chromosome 2	circular	88,746 bp	44.21%

The annotation of the *H. ammodendron* is summarized in Table 2. A total of 31 protein-coding genes, 18 tRNA, and three rRNA were predicted. Among protein-coding genes, 24 core genes usually present within angiosperms [21], and seven genes variably present were identified. The core genes consist of five ATP synthase genes (*atp1*, *atp4*, *atp6*, *atp8*, and *atp9*), nine NADH dehydrogenase genes (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, and *nad9*), four cytochrome c biogenesis genes (*ccmB*, *ccmC*, *ccmFC*, and *ccmFN*), three cytochrome c oxidase genes (*cox1*, *cox2*, and *cox3*), one membrane transport protein gene (*mttB*), one mature enzyme gene (*matR*), and one ubiquinol-cytochrome c reductase gene (*cob*). The non-core genes include one ribosomal large subunit gene (*rpl5*), five ribosomal small subunit genes (*rps3*, *rps4*, *rps7*, *rps12*, *rps13*), and one succinate dehydrogenase gene (*sdh4*). As for tRNA, three were in duplicate copy (*trnC-GCA*, *trnG-GCC*, *trnP-UGG*), while one was in triplicate copy (*trnfM-CAU*). Among the three rRNA, *rrn26* was present in the duplicate copy.

Table 2
H. ammodendron mitochondrial genome gene content information

Group of genes		Name of Genes
Core genes	ATP synthase	*atp1*, *atp4*, *atp6*, *atp8*, *atp9*
	NADH dehydrogenase	*nad1, nad2, nad3, nad4(×2), nad4L, nad5, nad6, nad7, nad9*
	Cytochrome c biogenesis	*ccmB*, *ccmC*, *ccmFC*, *ccmFN*
	Ubiquinol cytochrome c reductase	*cob*
	Cytochrome c oxidase	*cox1*, *cox2*, *cox3*
	Maturases	*matR*
	Transport membrane protein	*mttB*
Variable genes	Large subunit of ribosome	*rpl5*
	Small subunit of ribosome	*rps3*, *rps4*, *rps7*, *rps12*, *rps13*
	Succinate dehydrogenase	*sdh4*
rRNA genes	Ribosome RNA	*rrn5*, *rrn18*, *rrn26* (×2)
tRNA genes	Transfer RNA	*trnC-GCA(×2), trnD-GUC, trnE-UUC, trnFGAA, trnfM-CAU ×3 , trnG-GCC ×2 , trnHGUG, trnI-CAU, trnK-UUU, trnM-CAU, trnNGUU, trnP-UGG ×2 , trnQ-UUG, trnS-GCU, trnS-GGA, trnS-UGA, trnW-CCA, trnY-GUA*

Gene transfer between organelles

Putative gene transfer between organelles was uncovered by aligning the newly assembled chloroplast genome (Fig. 2A) with the mitochondrial genome of *H. ammodendron*. Homologous sequence fragments between the organelles are depicted in Fig. 2B. According to the sequence similarity analysis, 17 DNA fragments were identified as homologous between the mitochondrial and chloroplast genomes, with a total length of 5,233 bp, accounting for 2.49% of the mitochondrial genome (Table S3). Among them, the longest was 1,158 bp in length. By annotating these homologous sequences, eight complete genes were found on the 17 homologous fragments, including one protein-coding gene (*petG*) and seven tRNA genes (*trnD-GUC*, *trnH-GUG*, *trnM-CAU*, *trnN-GUU*, *trnP-UGG*, *trnS-GCU*, *trnW-CCA*).

Repeat sequences analysis

Microsatellites were screened for each set of chromosomes of *H. ammodendron* mitochondrial genome (Fig. 3A). For chromosome 1, a total of 42 simple sequence repeats (SSRs) were identified, comprising eight monomeric SSRs, eight dimeric SSRs, four trimeric SSRs, 17 tetrameric SSRs, four tetrameric SSRs, and one hexameric SSRs (Fig. 1A). Among SSRs, the repeat motifs $(A)_{10}$, $(GAAA)_3$ and $(CT)_5$ were found to be prevalent (Table S4). 17 tandem repeat sequences were identified, exhibiting a matching identity greater than 70% and lengths ranging from 14 to 31 bp (Fig. 3B, Table S5). Dispersed repeat sequences in chromosome 1 were also examined, detecting 467 repeat sequence pairs with a length equal to or greater than 30 bp. Among these, 234 pairs were palindromic repeats, 233 were forward repeats, and one was found to be a reverse repeat type. The longest observed palindromic repeat sequence spanned 123 bp, while the longest forward repeat sequence was 175 bp in length (Table S6).

In chromosome 2, 24 SSRs were detected (Fig. 3A). Among these SSRs, nine were monomeric, two were dimeric, three were trimeric, nine were tetrameric, and one was the pentameric type. No hexameric SSRs were found in this chromosome. The most prominent SSR motif was thymine $(T)_{10,}$ among others (Table S4). A total of 10 tandem repeat sequences with a matching identity greater than 78% and lengths ranging from 18 to 42 bp were identified (Fig. 3B, Table S5). Dispersed repeat sequences in chromosome 2 were also examined, detecting 148 pairs of repeat sequences with a length equal to or greater than 30 bp. Among these, 69 pairs were palindromic repeats, and 79 were forward repeats. No reverse or complement repeats were detected in both chromosomes. The longest observed palindromic repeat sequence spanned 56 bp, while the longest forward repeat sequence was 114 bp in length (Table S6).

RNA editing

RNA editing, a widespread phenomenon in land plants, entails modifications within the coding region of the transcript involving the addition, loss, or conversion of nucleotides [22]. The current study identified 337 RNA editing sites, all cytidine to uridin (C to U) conversion types. The predicted RNA editing sites in various genes are shown in Fig. 4. The *ccmB* and *ccmFN* genes exhibited the highest number (25) of RNA editing sites. Conversely, no RNA editing sites were observed in the *sdh4* gene. The amino acid transition exhibiting the highest abundance was from serine to leucine, comprising 19.58% (66 sites) of the observed transitions (Table S7).

Codon usage

The codon usage pattern of the *H. ammodendron* mitogenome was presented in Fig. 5. Codons exhibiting a relatively synonymous codon usage (RSCU) value greater than one was regarded as being favored by the corresponding amino acids. Apart from the start codon AUG and the tryptophan codon UGG, both of which had an RSCU value of 1, there were widespread codon usage preferences in the *H. ammodendron* mitochondrial protein-coding genes (Fig. 5). For example, alanine (Ala) showed a strong preference for

the codon GCU, with the highest RSCU value of 1.64. In addition, leucine (Leu) preferred the codon UUA, with an RSCU value of 1.61 (Table S8).

Phylogenetic analysis and collinearity analysis

The phylogenetic tree based on mitogenome DNA (Fig. 6) was constructed using a maximum likelihood approach, incorporating a diverse set of 30 species representing various orders (Table S9). The resulting tree topology aligns with the most recent classification the Angiosperm Phylogeny Group (APG) provided. *H. ammodendron* and the Amaranthaceae species Spinacia oleracea, Chenopodium quinoa, and Suaeda glauca were placed within the Caryophyllales order. Interestingly, *H. ammodendron* exhibited a close phylogenetic relationship with *Suaeda glauca*, another member of the Amaranthaceae family, widely distributed in the arid areas of northwest China [23].

To identify both conserved and unique mitochondrial regions among Caryophyllales species, we performed pairwise alignments using *Spinacia oleracea* (NC_035618.1), *Suaeda glauca* (NC_060419.1), *Alternanthera philoxeroides* (MN166292.1), *Beta macrocarpa* (NC_015994.1), *Silene latifolia* (NC_014487.1), *Mirabilis jalapa* (NC_056991.1), and *Fallopia aubertii* (MW664926.1) mitochondrial genomes (Fig. 7). The alignment results were provided in Fig. 7 and Table S10. Although collinear blocks were noticeable within Caryoplyllales species, their length is relatively short, indicating a non-conservative mitogenome structure. It is worth pointing out that the homologous block arrangement is inconsistent between *H. ammodendron* and its congeners, implying an evolutionary mitochondrial genome rearrangement occurrence. Additionally, some blank regions were observed in *H. ammodendron*, representing species-specific sequences lacking homology with other species.

Discussion

In land plants, mitochondrial genomes have undergone extensive genomic reorganization and gene arrangements, resulting in rapid structural evolution. Studying plant mitochondrial genomes offers a path into their structural organization, genetic diversity, and evolutionary dynamics, which have implications for understanding their functional role in biological processes such as plant growth, energy production, and respiration.

In the present study, we took advantage of both short- and long-read sequencing data to assemble the first mitochondrial genome of *H. ammodendron*. The assembly mitogenome length was 210,149 bp, which is about 50% shorter than *Suaeda glauca* (NC_060419.1, assembly size: 474,330 bp), and *Pereskia aculeate* (NC_067638.1, assembly size: 515,187 bp) but in similar range with *Mirabilis jalapa* (NC_056991.1, assembly size: 267,334 bp), all belonging to the same order: Caryophyllales. The observed genome size variation could be imputable to several factors, including repeat elements density, intragenic regions gain or loss, and the introgression of foreign DNA [24–26].

One of the complexities of the mitogenome is the variability in molecular conformation. Several configurations have been documented, including circular, linear, sigma-like, fragmented, and multi-

chromosomal structures [27]. Here we reported a multi-chromosomal mitogenome architecture that consists of two circular molecules. Similar findings were reported in other Caryophyllales species, such as *Fagopyrum esculentum* (10 circular chromosomes) [28], and *Fallopia multiflora* (two circular chromosomes) [29]. In the same plant order, a multi-chromosomal architecture with more than 50 chromosomes was also found for one of the symbolic plants, *Silene conica*, well known for its massive mitogenome spanning 11 Mb [14]. However, the close relatives *Suaeda glauca* (NC_060419.1) [18] and *Chenopodium quinoa* (NC_041093.1) [17] present a single molecule in a fully circularized conformation based on long reads data-based assembly. The origin and maintenance of the multi-ring structure in plant mitogenome are areas of ongoing research [30–35]. It is believed that mechanisms such as recombination and rearrangement events, possibly mediated by repeated sequences [36–38], could play a role in generating and maintaining this phenomenon. The balance between the recombination and replication of mitochondrial rings was proposed as a key aspect that could influence the occurrence of mitogenome multi-chromosome structure [39, 40]. Besides, gain or loss of chromosome sets is suggested to be one of the key biological events leading to the diversity of mitochondrial genome structures across different organisms [31], providing insights into the mechanisms driving the mitochondrial evolution.

Repeats elements such as microsatellites, tandem and dispersed repeats are prone to recombination leading to isomeric sequence sets in the mitogenome [14, 41, 42]. Therefore, the repeats dynamism could impact DNA maintenance and evolution of the plant mitochondrial genome [16].

Land plants' mitogenome globally exhibited many introns, some of which are split, requiring a separate transcription and subsequent trans-splicing of the resulting gene fragments [43]. Besides, incorporating foreign DNA influence could also influence the size mitogenome size variation. Our study detected sequence transfer between the chloroplast and the mitogenome, indicating possible horizontal or intracellular gene transfer [44].

RNA editing is a post-transcriptional mechanism in higher plant organelles, contributing to improved protein folding [45]. In angiosperm mitochondria, extensive cytidine to uridine RNA editing occurred, playing a crucial role in transcript expression [46]. Interestingly, all protein-coding genes analyzed from *H. ammomendron* mitogenome showed C- to-U conversion, which may affect the start or end positions of the coding genes. It is worth noting that proteins translated from edited RNAs could play a critical function in an environment change context, resulting in plant acclimatization and survival capabilities [47]. Indeed, *H. ammodendron* is well known for its remarkable tolerance to saline and dry environments [48, 49].

The maximum likelihood phylogenetic inference resulted in a tree with strong support values, placing *H. ammodendron* in the close relative taxon *Suaeda glauca*. Our tree topology is in perfect agreement with the recent results based on the Angiosperm Phylogeny Group (AGP III) classification [50], ITS [51], plastid markers [52], and the recent nuclear whole genome sequence data [10], indicating the reliability of the mitogenome-based tree.

Conclusions

In the present study, the mitochondrial genome of *H. ammodendron* was characterized. The mitogenome exhibited a multi-chromosomal conformation with two circular molecules with a total length of 210,149 bp. It contains 31 protein-coding genes, 18 tRNA and 3 rRNA. A set of 17 sequence fragments were found homologous between the mitochondrial and chloroplast genomes, spanning 5,233 bp, accounting for 2.49% of the mitogenome. RNA editing assessment revealed that the protein-coding genes were mainly subjected to cytidine-to-uridine conversion. The phylogenetic inference showed a close relationship with *S. glauca*. The mitogenome resource provided by this study laid a foundation for further investigations into the comparative evolution and functional role of the *H. ammodendron* mitogenome.

Methods

Assembly and annotation

Before the mitogenome assembly, short-reads data were trimmed using fastp v0.23.4 [53]. To assemble the *H. ammodendron* mitogenome, a hybrid approach encompassing both short and long-read sequencing data was implemented. Firstly, using the Illumina short-reads data, a draft mitogenome assembly was conducted with GetOrganelle v1.7.5 [19] with the following settings:-R 15 -k 21,45,65,85,105 -F embplant_pt. The resulting genome graph sketch was then visualized using Bandage v. 0.8.1 [20], manually inspected and curated for node junctions. Using BWA v0.7.17, we mapped PacBio HiFi data onto the initial assembly to address conflictual nodes and repetitive regions. [54]. In cases where multiple alternative connections exist at branching nodes, preference was given to connections supported by long reads. The obtained assembly was validated by mapping both short and long reads to ensure the connectivity and consistency of the assembly.

The assembly was annotated with GeSeq online tool (https://chlorobox.mpimp-golm.mpg.de/geseq.html#)[55] with the following parameters: *Arabidopsis thaliana* (NC_037304), *Nicotiana tabacum* (NC_006581), *Glycine max* (NC_020455), and *Suaeda glauca* (NC_060419.1) were set as reference mitogenomes; BLAT search for coding protein genes, rRNA and tRNA prediction. Additionally, tRNAscan-SE v2.0.7 [56] was used for tRNA prediction. The annotated mitogenome was visualized using Apollo v2.7.0 [57] for manual correction. The genome map was rendered using the Organellar Genome DRAW (OGDRAW) tool [58].

Repeat sequence analysis

The identification of repetitive sequences, including microsatellite sequence repeats (SSRs), tandem repeats, and dispersed repeats, was performed using MISA (https://webblast.ipk-gatersleben.de/misa/) [59], TRF () [60], and the REPuter (https://bibiserv.cebitec.uni-bielefeld.de/reputer/) [61] web servers respectively.

RNA editing and Codon usage bias analyses

The protein-coding sequences of the mitochondrial genome of *H. ammodendron* were extracted using the PhyloSuite toolbox [62]. Prediction of RNA editing sites was performed using the online tool PREPACT3 (http://www.prepact.de/) [63], while MEGA v7.0 software [64] was employed to calculate the relative synonymous codon usage (RSCU).

Collinearity analysis

Using the BLASTN v2.10.1 + tool, we conducted pairwise alignments to assess sequence similarity between *H. ammodendron* and its related species [65]. Homologous sequences exceeding 500 bp were retained as conserved collinear blocks for constructing the multiple synteny plot with the help of MCscanX implemented in TBTools v1.098746 [66].

Gene transfer analysis

To detect the putative gene transfer between chloroplast and mitochondrial genomes, we performed a *de novo* assembly and annotation of the chloroplast genome of our sample using GetOrganelle v1.7.5 [19] and GeSeq web server [55], respectively. Therefore, an alignment of organelles was performed using BLAST v2.10.1+ [65] with the following settings: E-value ≤ 1E-10, matching rate ≥ 70%, and matching length ≥ 40 bp.

Phylogenetic analysis

To infer the mitogenome-based phylogenetic tree, a set of 15 conserved genes (*atp1*, *atp4*, *atp6*, *atp8*, *ccmB*, *ccmC*, *ccmFC*, *ccmFN*, *cox2*, *matR*, *nad1*, *nad2*, *nad3*, *nad5*, and *nad6*) from 30 species belonging to four orders (Caryophyllales, Ericales, Santalales, and Zygopgyllales) (Table S10) were extracted from each mitogenome and aligned using MAFFT v7.505 [67]. Poorly aligned sequence regions were trimmed with trimAl v 1.4.1 [68]. The resulting data matrix was then concatenated with PhyloSuite [62], and the tree was constructed using IQ-TREE v2.0.3 [62]. The tree was rendered with the Interactive Tree of Life tool (iTOL) v5.0 [69].

Declarations

Ethics approval and consent to participate

Experimental research and studies on *H. ammodendron* comply with relevant institutional, national, and international guidelines and legislation.

Consent for publication

Not applicable

Data availability statement

The short reads data are available at NCBI SRA database under accession number SRR17127859. The PacBio dataset are available at NCBI SRA under the accession SRR17129371.

Competing of interests

The authors declare that they have no competing interest.

Funding

This study was supported by grants from the National Natural Science Foundation of China (82173066, 81974366), the Natural Science Foundation of Guangdong Province (2020A1515011408), the Shenzhen Science and Technology Innovation Commission Project (JCYJ20200109120208018).

Authors' contributions

L.Y. and T.X. designed the research. W.G. and Z. Z. participated in the material preparation. L.Y. and W.G., performed the genome assembly, and evolutionary analysis. L.Y. and T.X. wrote the manuscript.

Acknowledgements

We sincerely thank the experimental personnel and bioinformatics analysts at Wuhan Benagen Technology Company Limited (www.benagen.com) and MitoRun research group who participated in this project.

References

1. Su P, Cheng G, Yan Q, Liu X. Photosynthetic regulation of C4 desert plant *H. ammodendron* under drought stress. Plant Growth Regul. 2007;51:139–47.
2. *H. Bunge* (Chenopodiaceae). In: Flora of China Editorial Committee, ed. Flora of China. Vol. 5. St. Louis: Science Press, Beijing/Missouri Botanic Garden Press, 395–396. CABI Compendium. 2004;CABI Compe.
3. Yang WB, Feng W, Jia ZQ, Zhu YJ, Guo JY. Soil water threshold for the growth of *H. ammodendron* in the Ulan Buh desert in arid northwest China. South Afr J Bot. 2014;92:53–8.
4. Ma Q, Wang X, Chen F, Wei L, Zhang D, Jin H. Carbon sequestration of sand-fixing plantation of *H. ammodendron* in Shiyang River Basin: Storage, rate and potential. Glob Ecol Conserv. 2021;28:e01607.
5. Li J, Chang H, Liu T, Zhang C. The potential geographical distribution of *Haloxylon* across Central Asia under climate change in the 21st century. Agric For Meteorol. 2019;275:243–54.
6. Long Y, Zhang J, Tian X, Wu S, Zhang Q, Zhang J, et al. De novo assembly of the desert tree *H. ammodendron* (C. A. Mey.) based on RNA-Seq data provides insight into drought response, gene discovery and marker identification. BMC Genomics. 2014;15:1–11.
7. Gao HJ, Lü XP, Zhang L, Qiao Y, Zhao Q, Wang YP et al. Transcriptomic profiling and physiological analysis of *H. ammodendron* in response to osmotic stress. Int J Mol Sci. 2018;19.

8. Chen Y, Ma S, Zhang D, Wei B, Huang G, Zhang Y, et al. Diversification and historical demography of Haloxylon ammodendron in relation to Pleistocene climatic oscillations in northwestern China. PeerJ. 2022;10:1–27.
9. Suo Z, Jia Z, Lu Q, Pan B, Jin X, Xu G, Distinguishing H et al. *persicum* and *H. ammodendron* (*H. Bunge*, Amaranthaceae) using DNA Marker. AASRI Procedia. 2012;1:305–10.
10. Wang M, Zhang L, Tong S, Jiang D, Fu Z. Chromosome-level genome assembly of a xerophytic plant, *H. ammodendron*. DNA Res. 2022;29:1–9.
11. Chen Z, Zhao N, Li S, Grover CE, Nie H, Wendel JF, et al. Plant Mitochondrial Genome Evolution and Cytoplasmic Male Sterility. CRC Crit Rev Plant Sci. 2017;36:55–69.
12. Gray MW, Burger G, Lang BF. Mitochondrial Evolution. Sci (80-). 1999;283:1476–81.
13. Gray MW. Mitochondrial evolution. Cold Spring Harb Perspect Biol. 2012;4.
14. Sloan DB, Alverson AJ, Štorchová H, Palmer JD, Taylor DR. Extensive loss of translational genes in the structurally dynamic mitochondrial genome of the angiosperm Silene latifolia. BMC Evol Biol. 2010;10.
15. Smith DR, Keeling PJ. Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. Proc Natl Acad Sci. 2015;112:10177–84.
16. Wynn EL, Christensen AC. Repeats of unusual size in plant mitochondrial genomes: Identification, incidence and evolution. G3 Genes. Genomes Genet. 2019;9:549–59.
17. Maughan PJ, Chaney L, Lightfoot DJ, Cox BJ, Tester M, Jellen EN, et al. Mitochondrial and chloroplast genomes provide insights into the evolutionary origins of quinoa (*Chenopodium quinoa* Willd). Sci Rep. 2019;9:1–11.
18. Cheng Y, He X, Priyadarshani SVGN, Wang Y, Ye L, Shi C, et al. Assembly and comparative analysis of the complete mitochondrial genome of *Suaeda glauca*. BMC Genomics. 2021;22:1–15.
19. Jin J-J, Yu W-B, Yang J-B, Song Y, DePamphilis CW, Yi T-S, et al. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 2020;21:241.
20. Wick RR, Schultz MB, Zobel J, Holt KE, Bandage. Interactive visualization of de novo genome assemblies. Bioinformatics. 2015;31:3350–2.
21. Adams KL, Qiu Y, Stoutemyer M, Palmer JD. Punctuated evolution of mitochondrial gene content: High and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. Proc Natl Acad Sci. 2002;99:9905–12.
22. Jobson RW, Qiu YL. Did RNA editing in plant organellar genomes originate under natural selection or through genetic drift? Biol Direct. 2008;3:1–22.
23. Zhang S, Ni X, Arif M, Yuan Z, Li L, Li C. Salinity influences Cd accumulation and distribution characteristics in two contrasting halophytes, *Suaeda glauca* and *Limonium aureum*. Ecotoxicol Environ Saf. 2020;191 November 2019:110230.
24. Rice DW, Alverson AJ, Richardson AO, Young GJ, Sanchez-Puerta MV, Munzinger J, et al. Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm Amborella. Sci (80-).

2013;342:1468–73.

25. Zhang S, Wang J, He W, Kan S, Liao X, Jordan DR, et al. Variation in mitogenome structural conformation in wild and cultivated lineages of sorghum corresponds with domestication history and plastome evolution. BMC Plant Biol. 2023;23:1–17.
26. Mower JP. Variation in protein gene and intron content among land plant mitogenomes. Mitochondrion. 2020;53:203–13.
27. Gualberto JM, Newton KJ. Plant Mitochondrial Genomes: Dynamics and Mechanisms of Mutation. Annu Rev Plant Biol. 2017;68 February:225–52.
28. Logacheva MD, Schelkunov MI, Fesenko AN, Kasianov AS, Penin AA. Mitochondrial genome of *Fagopyrum esculentum* and the genetic diversity of extranuclear genomes in buckwheat. Plants. 2020;9.
29. Kim CK, Kim YK. The multipartite mitochondrial genome of *Fallopia multiflora* (Caryophyllales: Polygonaceae). Mitochondrial DNA Part B Resour. 2018;3:155–6.
30. Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, et al. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol. 2012;10:e1001241.
31. Wu Z, Cuthbert JM, Taylor DR, Sloan DB. The massive mitochondrial genome of the angiosperm *Silene noctiflora* is evolving by gain or loss of entire chromosomes. Proc Natl Acad Sci U S A. 2015;112:10185–91.
32. Bi C, Qu Y, Hou J, Wu K, Ye N, Yin T. Deciphering the Multi-Chromosomal Mitochondrial Genome of *Populus simonii*. Front Plant Sci. 2022;13:1–14.
33. Li X, Zhe M, Huang Y, Fan W, Yang J, Zhu A. The Evolution of Mitochondrial Genomes between Two *Cymbidium* Sister Species: Dozens of Circular Chromosomes and the Maintenance and Deterioration of Genome Synteny. Genes (Basel). 2023;14:864.
34. Yu R, Sun C, Zhong Y, Liu Y, Sanchez-Puerta MV, Mower JP, et al. The minicircular and extremely heteroplasmic mitogenome of the holoparasitic plant *Rhopalocnemis phalloides*. Curr Biol. 2022;32:470–479e5.
35. Szandar K, Krawczyk K, Myszczyński K, Ślipiko M, Sawicki J, Szczecińska M. Breaking the limits - multichromosomal structure of an early eudicot *Pulsatilla patens* mitogenome reveals extensive RNA-editing, longest repeats and chloroplast derived regions among sequenced land plant mitogenomes. BMC Plant Biol. 2022;22:109.
36. Yang H, Ni Y, Zhang X, Li J, Chen H, Liu C. The mitochondrial genomes of *Panax notoginseng* reveal recombination mediated by repeats associated with DNA replication. Int J Biol Macromol. 2023;:126359.
37. Jiang M, Ni Y, Li J, Liu C. Characterisation of the complete mitochondrial genome of *Taraxacum mongolicum* revealed five repeat-mediated recombinations. Plant Cell Rep. 2023;42:775–89.
38. Sloan DB. One ring to rule them all? Genome sequencing provides new insights into the master circle model of plant mitochondrial DNA structure. New Phytol. 2013;200:978–85.

39. Kozik A, Rowan BA, Lavelle D, Berke L, Schranz ME, Michelmore RW, et al. The alternative reality of plant mitochondrial DNA: One ring does not rule them all. PLOS Genet. 2019;15:e1008373.
40. Morley SA, Ahmad N, Nielsen BL. Plant Organelle Genome Replication Plants. 2019;8:358.
41. Folkerts O, Hanson MR. Three copies of a single recombination repeat occur on the 443 kb mastercircle of the *Petunia hybrida* 3704 mitochondrial genome. Nucleic Acids Res. 1989;17:7345–57.
42. Siculella L, Damiano F, Cortese MR, Dassisti E, Rainaldi G, Gallerani R, et al. Gene content and organization of the oat mitochondrial genome. Theor Appl Genet. 2001;103:359–65.
43. Malek O, Brennicke A, Knoop V. Evolution of trans-splicing plant mitochondrial introns in pre-Permian times. Proc Natl Acad Sci U S A. 1997;94:553–8.
44. Archibald JM, Richards TA. Gene transfer: Anything goes in plant mitochondria. BMC Biol. 2010;8:8–10.
45. Hao W, Liu G, Wang W, Shen W, Zhao Y, Sun J et al. RNA Editing and Its Roles in Plant Organelles. Front Genet. 2021;12 September.
46. Knoop V. C-to-U and U-to-C: RNA editing in plant organelles and beyond. J Exp Bot. 2023;74:2273–94.
47. Gommans WM, Mullen SP, Maas S. RNA editing: A driving force for adaptive evolution? BioEssays. 2009;31:1137–45.
48. Hu D, Lv G, Qie Y, Wang H, Yang F, Jiang L. Response of morphological characters and photosynthetic characteristics of *H. ammodendron* to water and salt stress. Sustain. 2021;13:1–12.
49. He XH, Si JH, Zhou DM, Wang CL, Zhao CY, Jia B, et al. Leaf chlorophyll parameters and photosynthetic characteristic variations with stand age in a typical desert species (*H. ammodendron*). Front Plant Sci. 2022;13:1–13.
50. Bremer B, Bremer K, Chase MW, Fay MF, Reveal JL, Bailey LH, et al. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot J Linn Soc. 2009;161:105–21.
51. Pyankov VI, Artyusheva EG, Edwards GE, Black CC, Soltis PS. Phylogenetic analysis of tribe salsoleae (Chenopodiaceae) based on ribosomal its sequences: Implications for the evolutIon of photosynthesis types. Am J Bot. 2001;88:1189–98.
52. Akhani H, Edwards G, Roalson EH. Diversification of the Old World Salsoleae s. l. (Chenopodiaceae): Molecular Phylogenetic Analysis of Nuclear and Chloroplast Data Sets and a Revised Classification Author (s): Hossein Akhani, Gerald Edwards, and Eric H. Roalson DIVERSIFICATION O. Int J Plant Sci. 2007;168:931–56.
53. Chen S, Zhou Y, Chen Y, Gu J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–90.
54. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 2010;26:589–95.

55. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, et al. GeSeq - Versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 2017;45:W6–11.
56. Lowe TM, Eddy SR. tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. Nucleic Acids Res. 1997;25:0955–64.
57. Dunn NA, Unni DR, Diesh C, Munoz-Torres M, Harris NL, Yao E, et al. Apollo: Democratizing genome annotation. PLoS Comput Biol. 2019;15:1–14.
58. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 2019;47:W59–64.
59. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. Bioinformatics. 2017;33:2583–5.
60. Benson G. Tandem repeats finder: A program to analyze DNA sequences. Nucleic Acids Res. 1999;27:573–80.
61. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. REPuter: The manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 2001;29:4633–42.
62. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, et al. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol Ecol Resour. 2020;20:348–55.
63. Lenz H, Rüdinger M, Volkmar U, Fischer S, Herres S, Grewe F, et al. Introducing the plant RNA editing prediction and analysis computer tool PREPACT and an update on RNA editing site nomenclature. Curr Genet. 2010;56:189–201.
64. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33:1870–4.
65. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
66. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant. 2020;13:1194–202.
67. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
68. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3.
69. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49:W293–6.

Figures

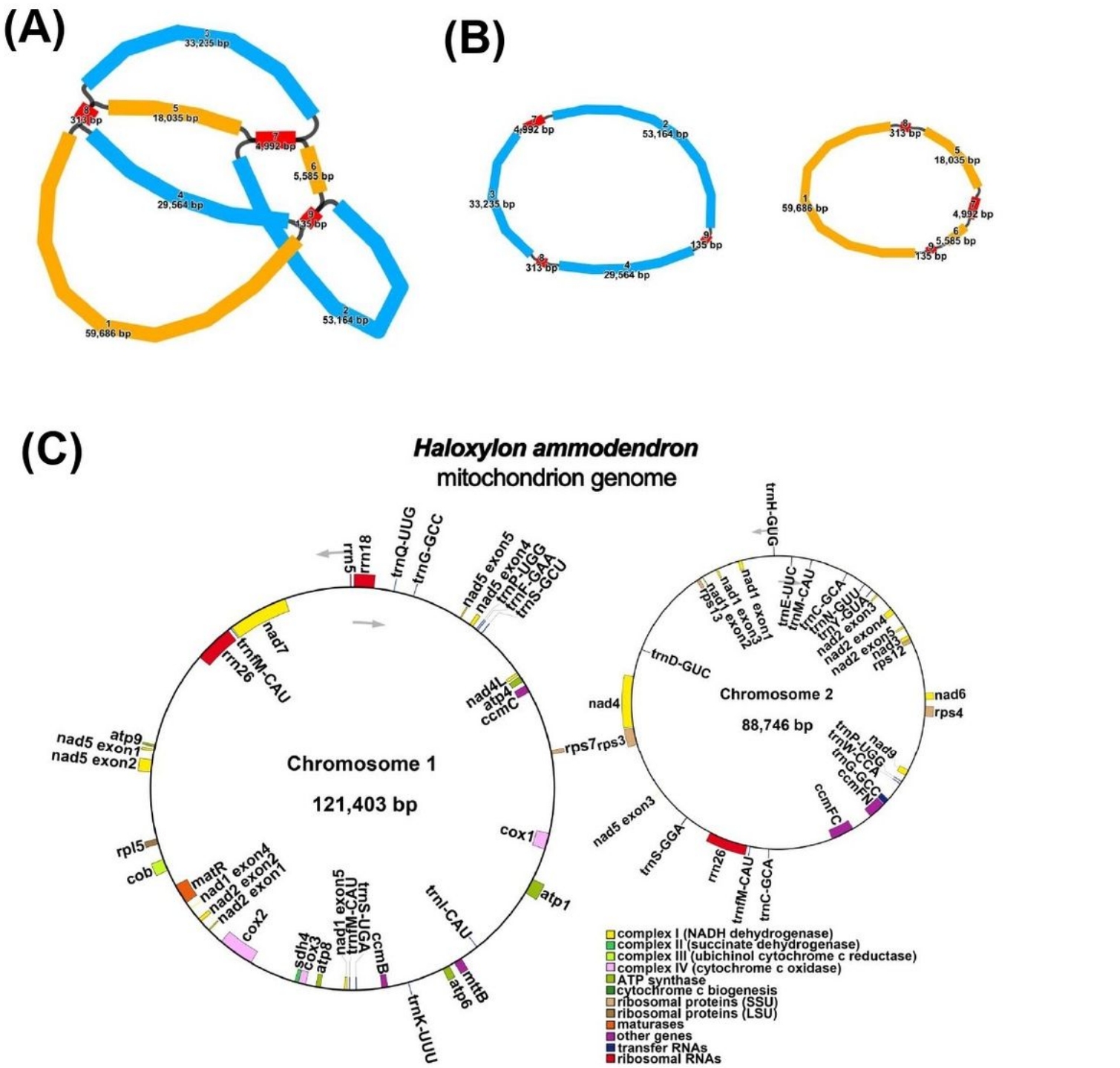


Figure 1

The mitochondrial genome of *H.ammodendron*. (A) Mitochondrial assembly graph view from Bandage tool showing nine nodes (or contigs). (B) Separated view of two sets of contigs network following an assembly path resolution. The first path grouped contigs 2, 9, 4, 8, 3, and 7, while the second path encompassed contigs 1, 8, 5, 7, 6, and 9. The black lines connecting two nodes indicate overlapping regions between the contigs. Red nodes represent potential repetitive sequences. (C) Mitogenome map of

the two fully circularized sequences. Genes inside the outer circle are clockwise, whereas those outside the outer circle are counterclockwise.

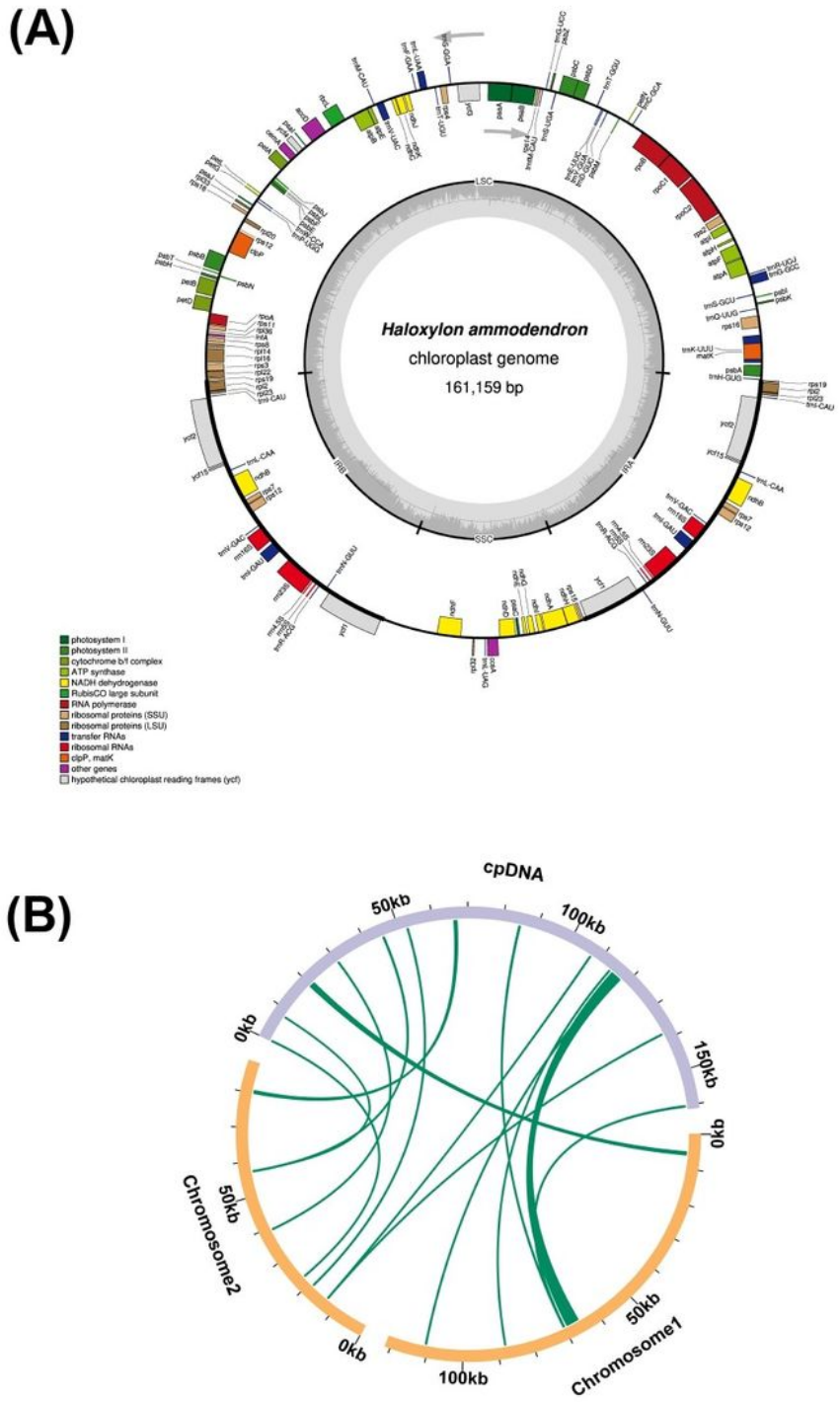


Figure 2

***Haloxylon ammodendron*plastome overview and sequence similarity.** (A) Depiction of the *H. ammodendron* plastome, providing a visual representation of its organization and key features. (B)

Illustration highlighting sequence similarity between the mitochondrial and chloroplast organelles. Yellow arcs denote regions of the mitochondrial genome, while the purple arc signifies the chloroplast genome. The interconnecting green lines emphasizehomologous segments shared between the two genomes.

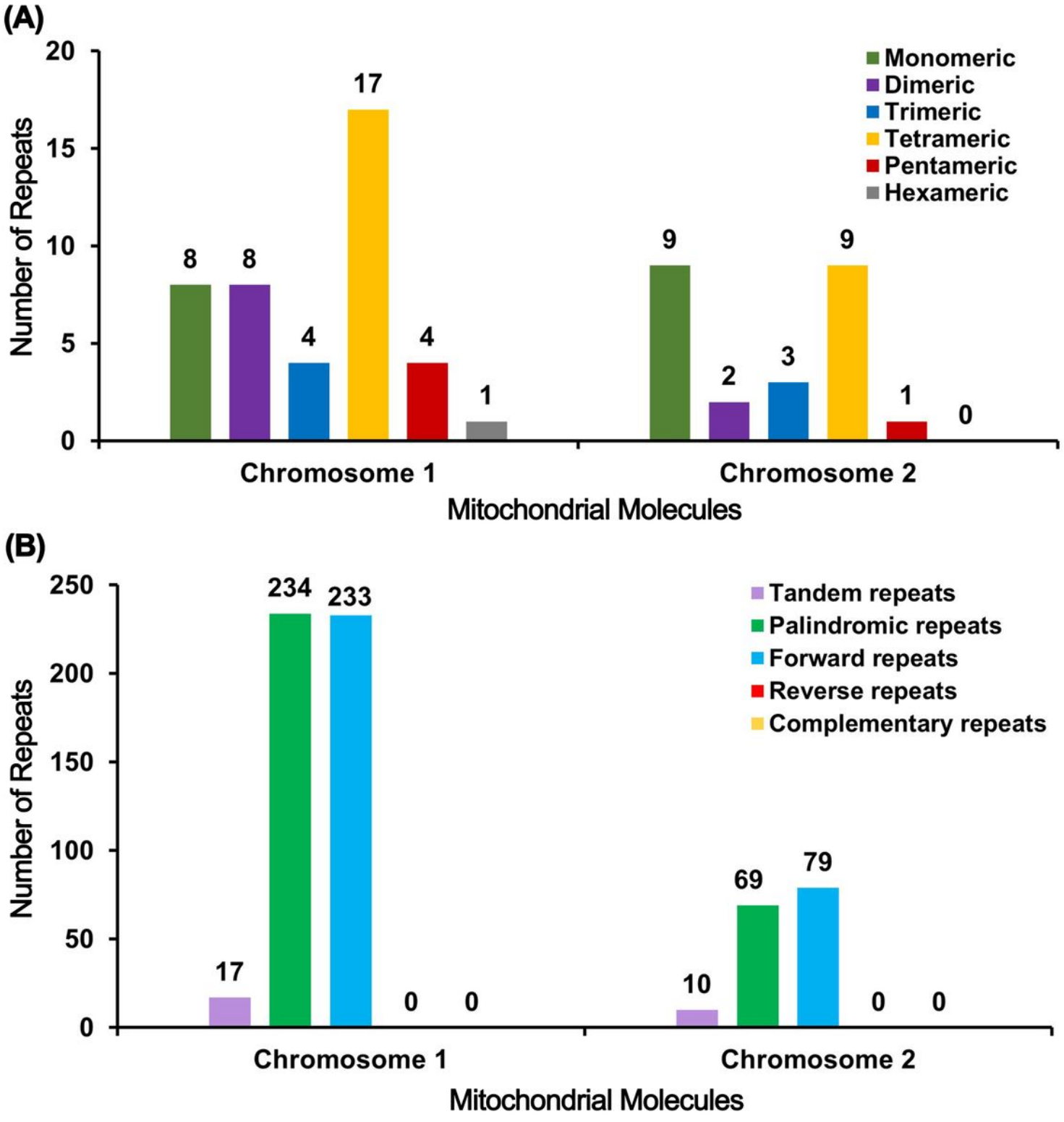


Figure 3

Repeat patterns in the *H. ammodendron* mitochondrial genome.

(A) Chart showing the count of Simple Sequence Repeats (SSRs) distributed in chromosome 1 and chromosome 2. (B) Visualization of the number of tandem and dispersed repeats found in both chromosomes, offering insight into the genome's complexity and evolutionary history.

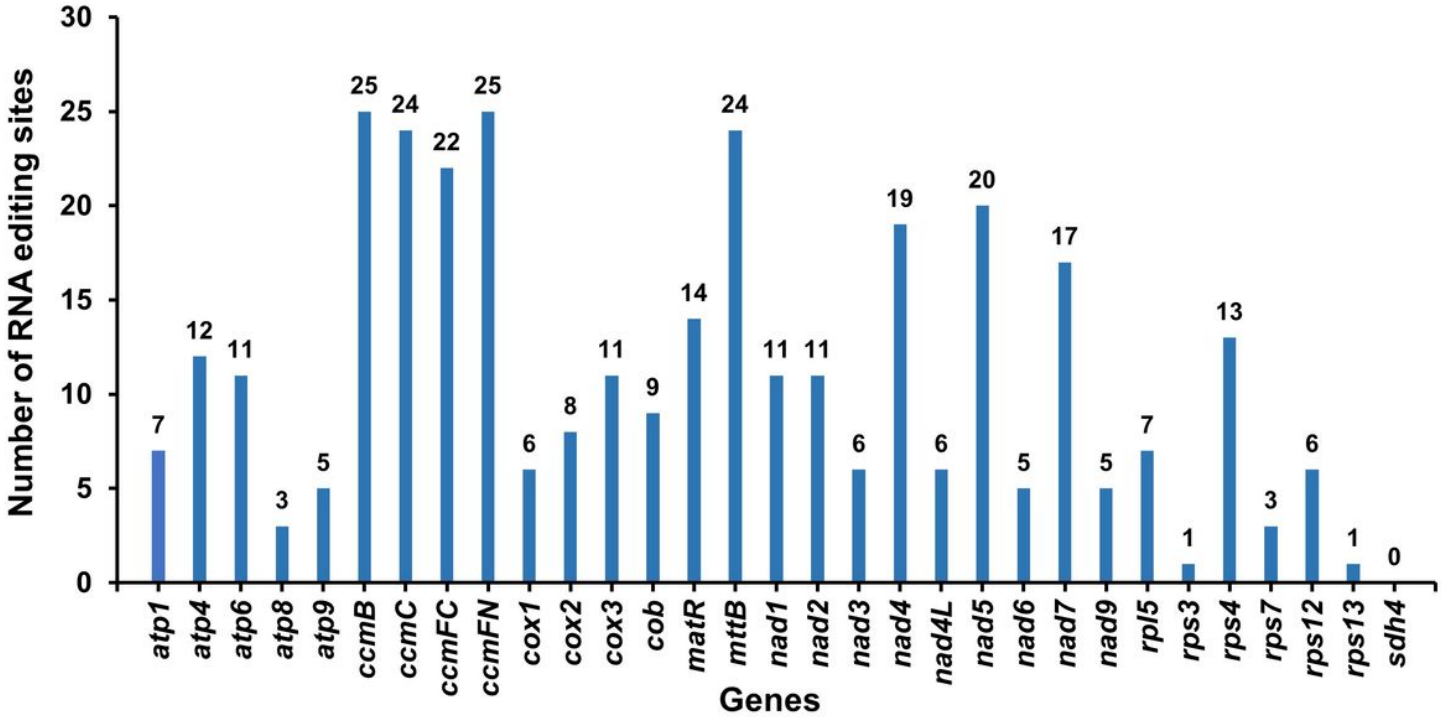


Figure 4

RNA editing sites in *H. ammodendron* mitochondrial genes. This figure presents a comprehensive count of RNA editing sites discovered within 31 mitochondrial protein-coding genes of *H. ammodendron*, shedding light on post-transcriptional modifications and their potential implications.

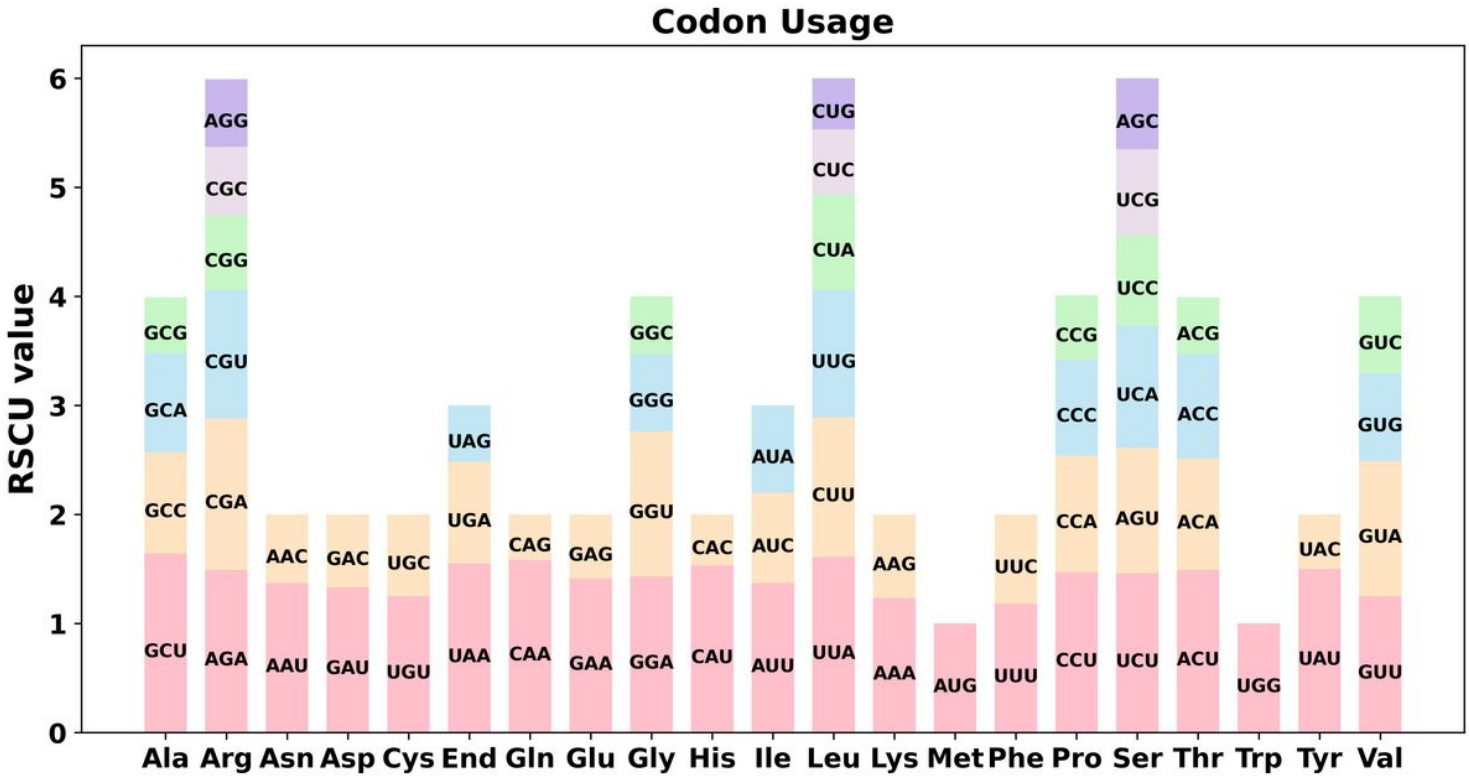


Figure 5

Codon usage preferences in *H. ammodendron* 's mitochondrial genome. Graphical representation illustrating the Relative Synonymous Codon Usage (RSCU) values. This showcases the preferences of certain codons encoding the same amino acid, hinting at evolutionary pressures and translational optimization.

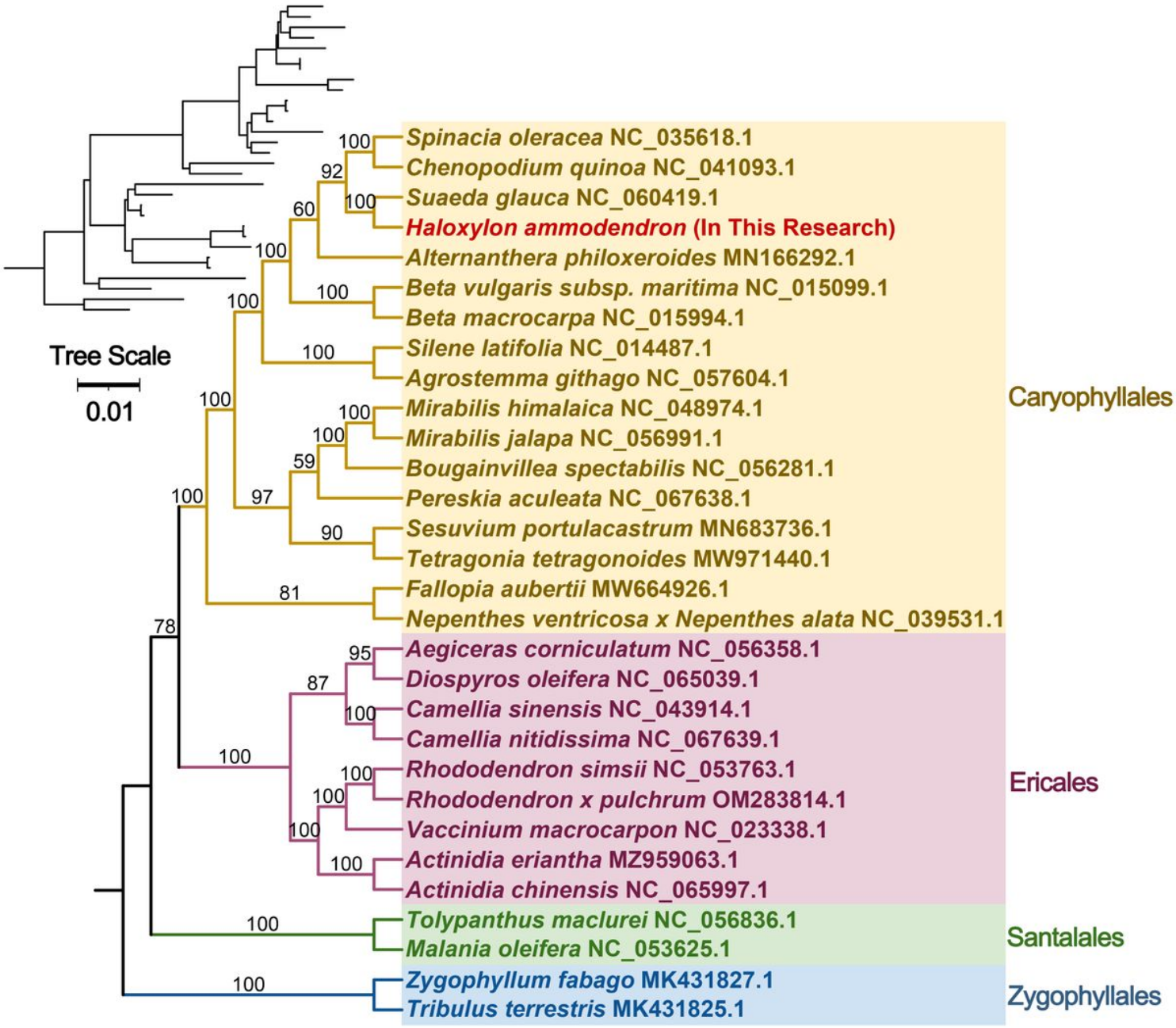


Figure 6

Phylogenetic placement of *H. ammodendron* within Caryophyllales. A Maximum Likelihoodphylogenetic tree that delineates the evolutionary relationship of *H. ammodendron* with other species in the Caryophyllales order. Node labels display support values, with maximum-likelihood bootstrap values offering a measure of the tree's reliability.

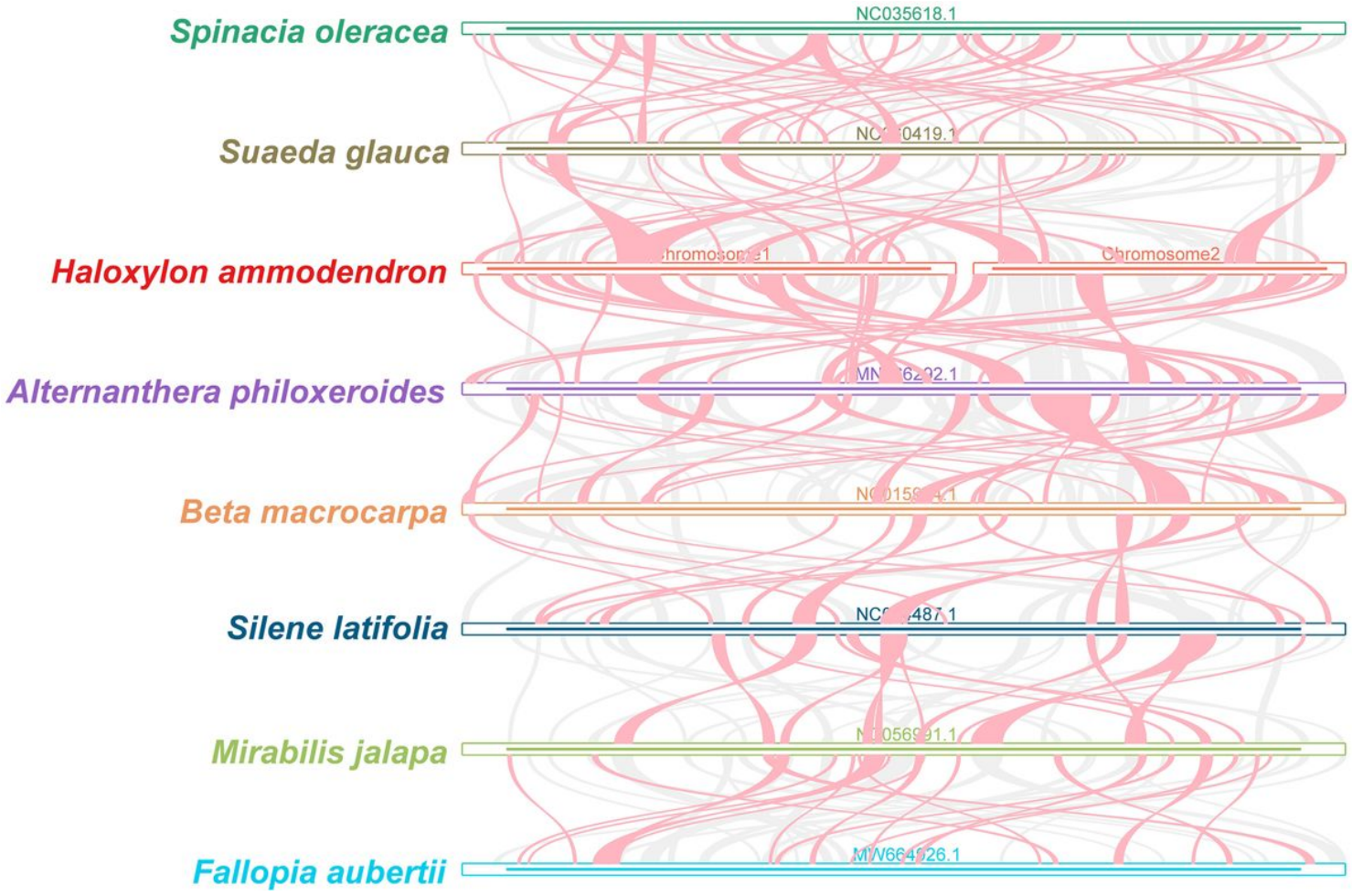


Figure 7

Comparative collinearity in Caryophyllales mitogenomes. A landscape view showcasing the genome collinearity between *H. ammodendron* and itsCaryophyllales counterparts. The diagram helps in understanding genomic rearrangements, evolutionary events, and conservation levels among these species.

Supplementary Files

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

- SupplementaryTables.xlsx