

Mitogenome Architecture and Evolutionary Insights in *Origanum vulgare*: Comparative Analysis of Structural Diversity and Repeat-Mediated Recombination

Meiling Qin

Affiliated Hospital of Jiangsu University

Shaoshuai Yu

Affiliated People's Hospital of Jiangsu University

Ikram Outissint

Jiangsu University

Min Tang

Jiangsu University

Xun Gong

1000012836@ujs.edu.cn

Affiliated Hospital of Jiangsu University

Research Article

Keywords: *Origanum vulgare*, mitogenome, repeat-mediated recombination, RNA editing, phylogeny

Posted Date: July 24th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-7070486/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Mitogenome Architecture and Evolutionary Insights in *Origanum vulgare*: Comparative Analysis of Structural Diversity and Repeat-Mediated Recombination

Meiling Qin^{1,2}, Shaoshuai Yu³, Ikram Outissint², Min Tang², Xun Gong^{1*}

1. Department of Rheumatology & Immunology, Affiliated Hospital of Jiangsu University, Zhenjiang 212001, Jiangsu, China

2. School of Life Sciences, Jiangsu University, Zhenjiang 212013, Jiangsu, China

3. Department of Pharmacy, Affiliated People's Hospital of Jiangsu University, Zhenjiang 212001, Jiangsu, China

* Corresponding author: Xun Gong (1000012836@ujs.edu.cn)

Number of Pages: 44

Number of Figures: 11

Number of Tables: 2

ABSTRACT

Origanum vulgare is a perennial herb known for its antimicrobial, antioxidant, and anti-inflammatory properties, valued in both traditional medicine and modern phytopharmacology. Despite substantial research on its chemical and pharmacological properties, genomic studies have been limited. In this study, we present the first comprehensive assembly and annotation of the mitochondrial genome (mitogenome) of *O. vulgare* using HiFi long-read sequencing technology. The assembled genome spans 497,243 base pairs and includes 57 protein-coding genes (PCGs), 14 transfer RNA genes, and 4 ribosomal RNA genes. Additionally, we identified 9 simple sequence repeats, 26 tandem repeats, and 607 dispersed repeats, alongside 395 RNA editing sites, predominantly located at the first and second codon positions, suggesting a significant role in gene expression regulation. Phylogenetic analysis of mitochondrial PCGs from 25 representative species confirmed *O. vulgare*'s close evolutionary relationship with other species within the Lamiales order, highlighting significant genomic rearrangements and adaptive evolution. We also completed the assembly, annotation, and comprehensive analysis of the chloroplast genome of *O. vulgare*, providing further insights into its organellar genomes. Our findings further indicate that the mitogenome exhibits high levels of structural variability, which may contribute to the plant's ability to adapt to diverse environmental conditions. The identification of RNA editing events suggests a key role in post-transcriptional regulation, influencing protein diversity and potentially enhancing stress tolerance. This research establishes a genomic foundation for *O. vulgare*, advancing our understanding of its mitogenome structure, molecular evolution, and adaptive mechanisms, with important implications for phytopharmacological research and biotechnological applications in plant breeding.

Key words: *Origanum vulgare*; mitogenome; repeat-mediated recombination; RNA editing; phylogeny

INTRODUCTION

Origanum vulgare (common oregano) belongs to the Lamiaceae family, order Lamiales, and is a perennial herb native to the mediterranean region, Europe, and parts of Asia. It is characterized by high concentrations of phenolic compounds, primarily carvacrol (60–80%) and thymol (5–10%), alongside rosmarinic acid, flavonoids, and terpenoids [1, 2]. Recent pharmacological investigations have revealed that oregano extracts exhibit remarkable antimicrobial activity against a broad spectrum of pathogens, including methicillin-susceptible and methicillin-resistant Staphylococci (MSS, MRS) with minimum inhibitory concentrations as low as 0.06–0.125% (v/v), *Candida albicans* (MIC: 0.4 mg/mL) and food-spoilage fungi [1, 3]. The mechanism of action involves disruption of microbial cell membranes and inhibition of biofilm formation, making it particularly effective against drug-resistant strains [4]. Beyond its antimicrobial properties, oregano demonstrates significant antioxidant capacity, showing an Oxygen Radical Absorbance Capacity (ORAC) value of 183,141 $\mu\text{mol TE}/100\text{g}$, significantly higher than many spices (e.g., apple: 520 $\mu\text{mol TE}/100\text{g}$; agave: 1247 $\mu\text{mol TE}/100\text{g}$; tarragon: 15542 $\mu\text{mol TE}/100\text{g}$) [5], meanwhile its anti-inflammatory properties are evidenced by significant antinociception and activated opioid receptor-NO-cGMP-K⁺channel pathway [6]. What's more, current research has expanded into potential anticancer applications, with carvacrol showing dose-dependent apoptosis induction in human colon and breast cancer cell lines [7]. Agricultural applications include its use as a natural preservative, where 0.1% oregano essential oil extends the shelf-life of various food products by inhibiting spoilage microorganisms. Recent advances in formulation technology, particularly nanoencapsulation in lipid carriers, have significantly improved the bioavailability and stability of its active components. Though further studies are needed to fully elucidate its mechanisms of action and optimize its clinical applications, the growing body of research underscores oregano's potential as a source of novel therapeutics.

Extensive research has been conducted on the chemical constituents and pharmacological properties of *O. vulgare*, while genomic investigations, particularly

phylogenetic and mitogenomic analyses, remain limited. This underscores the critical need for mitogenome studies, which serve as powerful molecular markers for elucidating evolutionary relationships and phylogenetic history. As essential organelles in plant cells, mitochondria play a pivotal role in cellular energy metabolism, primarily through oxidative phosphorylation to generate ATP. Distinct from their animal counterparts, plant mitochondria exhibit several unique genomic features, including substantially larger genome sizes, more intricate structural organization, and elevated recombination frequencies [8]. Beyond its central role in cellular bioenergetics, the mitochondrion serves as a crucial regulatory hub for multiple metabolic processes, including lipid biosynthesis, amino acid homeostasis, and nucleotide metabolism. Furthermore, these organelles are integral to programmed cell death pathways and function as key modulators of intracellular calcium dynamics [9-12]. Therefore, it is crucial to carry out the study for plant mitogenomes, which help reveal how plants adapted to different environmental conditions and what genetic changes they have made during evolution. Then, these genetic changes can be beneficial to develop new variety breeding with specific traits [13-15].

In this study, we used DNA barcoding as a novel approach for precise species identification in *O. vulgare*. Then, we present the first complete assembly and annotation of both organellar genomes (mitochondrial and chloroplast), including detailed characterization of repetitive elements and codon usage patterns, significantly advancing genomic research in the Lamiales order. Furthermore, we performed systematic prediction of RNA-editing events in mitochondrial protein-coding genes (PCGs), revealing their widespread impact on protein-coding transcripts, non-coding regions, structural RNAs, and intronic sequences through post-transcriptional modification [16]. Additionally, the mitochondrial plastid sequences (MTPTs) as well as the regional boundaries of chloroplasts were identified. Furthermore, phylogenetic reconstruction utilizing mitochondrial PCGs from 25 representative species elucidated evolutionary relationships among *O. vulgare* and related taxa across five key angiosperm orders (Gentianales, Solanales, Lamiales, Asterales, and Apiales). Complementary analyses of nucleotide diversity and selection pressures revealed

patterns of molecular variation and evolutionary constraints acting on protein-coding sequences. Comparative genomics identified conserved homologous collinear blocks, providing new understanding of syntenic relationships and structural evolution within Lamiales. These comprehensive findings not only characterize the unique genomic architecture of *O. vulgare*'s mitochondrion but also establish a crucial foundation for exploring the species' genetic diversity and functional adaptations.

MATERIALS AND METHODS

Collection of Plant Materials, Library Preparation, and Sequencing

Fresh *O. vulgare* leaf samples were collected from Jiangsu University in Zhenjiang Province, China (geographical coordinates: 32.11°N, 119.35°E). To ensure taxonomic accuracy, the specimens were authenticated by qualified botanists prior to experimental procedures. Following collection, the leaves were carefully washed with diethylpyrocarbonate (DEPC)-treated water to eliminate contaminants and then flash-frozen at -80°C for subsequent analysis. Genomic DNA was extracted using a cetyltrimethylammonium bromide (CTAB)-based method, specifically adapted for high-molecular-weight DNA recovery. The quality of the isolated DNA was assessed via 0.75% agarose gel electrophoresis to confirm structural integrity. Additionally, purity and concentration were evaluated using two complementary approaches: ultraviolet spectrophotometric analysis (NanoDrop One, Thermo Fisher Scientific) and fluorometric quantification (Qubit 3.0 Fluorometer, Life Technologies). For long-read sequencing, DNA samples meeting stringent quality criteria were processed with the SMRTbell Express Template Preparation Kit 2.0 (Pacific Biosciences). The resulting libraries were sequenced on a PacBio Sequel II platform to generate high-fidelity reads (Figure S1). The following analyses were carried out in accordance with Figure S2.

Species Authentication and Molecular Validation

To confirm taxonomic identity, the collected samples were subjected to morphological analysis by comparison with a voucher specimen (accession no. KUN1566203)

archived at the National Plant Specimen Resource Center of China Digital Herbarium, housed at the Kunming Institute of Botany, Chinese Academy of Sciences (Figure 1A). For further molecular verification, the core plant barcode region *rbcL* was amplified from genomic DNA using standardized primers (Table S1). The primer sequences (*rbcL*-F and *rbcL*-R) and PCR amplification parameters followed the protocol described in [17]. Following amplification, the PCR products were purified and subjected to bidirectional Sanger sequencing (Sangon Biotech, Shanghai). The obtained sequences were then compared against the Barcode of Life Data Systems (BOLD) database (v4, <https://v4.boldsystems.org/>) to assess taxonomic congruence [18]. Finally, species identification was determined based on alignment scores, percentage sequence similarity, and statistical significance (E-values), ensuring high-confidence classification.

Mitogenome Assembly and Annotation

The mitogenome of *O.vulgare* was assembled through a two-step process. In the primary phase, putative mitochondrial sequences were identified by conducting a BLASTn (v2.13.0+) alignment of PacBio HiFi reads against reference mitogenomes from three related Lamiaceae species: *Lavandula angustifolia* (NC_082307.1), *Platostoma chinense* (NC_069573.1), and *Salvia rosmarinus* (PP992923.1) [19]. This screening employed stringent parameters (-evalue 1e-5 -outfmt 6 -max_hsps 10 -word_size 7 -task BLASTn-short) to ensure specificity. Subsequent genome assembly was executed using PMAT (v1.5.3) with optimized parameters for mitochondrial reconstruction (autoMito -st hifi -g 700m -m -tp all -cpu 64), yielding the master structure [20] and a complete circularized mitogenome was obtained after optimization in Bandage [21]. For genome annotation, we employed a dual-platform strategy: initial gene prediction was performed using the PMGA web server (<http://47.96.249.172:16084/index.html>), which incorporated comparative analysis against 29 reference mitogenomes [22], followed by manual curation of annotation artifacts through GeSeq web server [23]. The final mitogenome map was generated using OGDRAW (v1.3.1) to visualize gene organization and structural features [24].

Repeat Sequence Analysis and Codon Usage Bias

Due to their compact nature compared to nuclear genomes, organelle genomes contain simpler repeat architectures. We systematically categorized repetitive elements into three classes: (1) simple sequence repeats (SSRs), (2) tandem duplications, and (3) interspersed repeats. SSR identification was conducted using MISA-web [25] with the following minimum repeat unit thresholds: 10 for mononucleotide, 6 for dinucleotide, and 5 for tri-, tetra-, penta-, and hexanucleotide motifs, with a maximum length cutoff of 1000 bp. Tandem repeats were detected via Tandem Repeats Finder (TRF v4.09) [26] using optimized parameters (2 7 7 80 10 50 2000 -f -d -m). Dispersed repeats were analyzed using REPuter tool (<https://bibiserv.cebitec.uni-bielefeld.de/reputer/>) [27] with Minimal Repeat Size set as 30 and Maxumum Computed Repeats set as 1000. All repeat elements were graphically represented using Circos (v0.69-8) [28].

Codon usage bias, reflecting the preferential use of specific synonymous codons, provides insights into both evolutionary selection pressures and gene expression optimization [29]. In addition, Genes employing tRNA-abundant codons demonstrate enhanced translational efficiency, suggesting their potential functional importance in cellular processes [30]. We quantified codon preference using relative synonymous codon usage (RSCU) values computed by Phylosuit (v1.2.3) [31], with results visualized using the “ggplot2” package in R. This analysis offers valuable perspectives on the evolutionary dynamics and expression patterns of organelle genes.

Assembly of *cpgenome* and Mitochondrial-Plastid DNA Transfer Analysis

The *cpgenome* of *O.vulgare* was assembled with Flye (v2.9.4-b1799) based on reads screened by BLASTn (v2.13.0+) [32]. The genome circle was annotated and visualized with the CPGAVAS2 web server (<http://47.96.249.172:16019/analyzer/annotate>) [33]. The Sequence identity plot were present with mVISTA website (<https://genome.lbl.gov/vista/mvista/submit.shtml>) [34-37].

Interorganellar gene transfer represents a well-documented phenomenon in higher plants, with approximately 5-10% of mitochondrial genomic content demonstrating plastid origin across various species [38]. Notably, certain repetitive elements within mitogenomes may originate from recurrent chloroplast DNA insertions. To investigate these mitochondrial-plastid transfer sequences (MTPTs), we performed comparative genomic analysis using BLASTn (v2.13.0), employing the mitogenome as the reference database and the plastome as query sequences under stringent parameters (e-value cutoff: 1e-6). This approach provides critical insights into horizontal gene transfer mechanisms and their evolutionary implications in plant organellar genomes. The identified MTPTs were subsequently visualized using TBtools (v2.010) [39].

RNA Editing Site Prediction and tRNA Structural Analysis

RNA editing represents a critical post-transcriptional modification mechanism in plant mitochondria, involving nucleotide alterations that influence gene expression and protein functionality [40]. To identify potential RNA editing sites, we employed the PREPACT3 online platform (<http://www.prepact.de/prepact-main.php>), accessed on 25 May 2025 [41]. The analysis utilized *Nicotiana tabacum* mitogenome as the reference sequence, with sequence alignments performed using a stringent BLASTX E-value threshold of 0.001 to ensure prediction accuracy. The RNA-editing site prediction results were summarized and visualized by “ggplot2” package, then the alignment results were visualized by Jalview 2.11.4.1 [42].

Transfer RNAs (tRNAs) serve as essential molecular adaptors in protein synthesis and participate in diverse cellular processes. Comprehensive characterization of tRNA structures provides critical insights into their role in genetic information flow. To predict tRNA secondary structures, we employed tRNAscan-SE 2.0, which enables accurate reconstruction of cloverleaf conformations based on sequence features [43].

Selective pressure Analysis and Nucleotide diversity Analysis

Natural selection acts as a fundamental evolutionary force that shapes genetic adaptation by differentially favoring advantageous mutations. To quantify selection

pressure, we employed the nonsynonymous-to-synonymous substitution rate ratio ($\omega = K_a/K_s$), where K_a represents nonsynonymous mutations that alter amino acid sequences and may affect protein function and K_s reflects synonymous mutations that do not change amino acid sequences and are generally considered evolutionarily neutral. This ratio serves as a powerful indicator of molecular evolution ($\omega > 1$ suggests positive selection: adaptive evolution, $\omega = 1$ indicates neutral evolution and $\omega < 1$ implies purifying selection: functional constraint). The analysis was performed using KaKs_Calculator 2.0 [44].

Nucleotide diversity (π) serves as a fundamental metric for quantifying genetic variation within populations, providing critical insights into evolutionary dynamics and adaptive potential. Higher π values indicate greater genetic heterogeneity within the population. Elevated diversity suggests enhanced adaptive capacity and evolutionary potential and reduced diversity may reflect recent bottlenecks or strong selective sweeps. The π values were calculated using DnaSP software [45].

Phylogenetic and Collinear Analysis of *O.vulgare* Mitogenome

24 species from five orders (Apiales, Asterales, Solanales, Lamiales and Gentianales) closely related to *O.vulgare* were selected, and *Panax quinquefolius* from Apiales were set as the outgroup for phylogenetic analysis (Table S11). All genomic sequences in GenBank format were processed using Phylosuit v1.2.3 (v1.2.3) [31]. Then MAFFT (v7.313) was used for sequence alignment [46], followed by maximum likelihood phylogenetic reconstruction with IQ-TREE2 (v2.1.4) [47]. Final iTOL were used for tree visualization and optimization [48].

Four representative Lamiales mitogenomes were selected as references for colinear comparison. Based on BLASTn (v2.13.0; e-value $\leq 1e-6$), we identified homologous sequences and in which >500bp were considered as conserved syntenic blocks. The syntenic blocks was visualized with NGenomeSyn (v1.41), Manve [49, 50] and MUMmer 3.0 [51].

RESULTS

Genetic Identification of Collected Specimens

The PCR amplification of the target *rbcL* and *matK* genes using the designed primer pair yielded specific products matching the expected fragment size (Table S1, Figure 1B). Subsequent sequence analysis revealed exceptionally high homology with *O. vulgare*, as demonstrated by a perfect match (100% similarity, score=1300) in the *rbcL* database comparison (Figure S3). Similarly, the *matK* gene product showed near-complete sequence identity (99.92% similarity, score=1298) with *O. vulgare* reference sequences (Figure S4), further confirming the taxonomic identity of the collected specimens. These molecular results, showing consistent high homology across two distinct genetic markers, provide robust validation of the sample identification and confirm the specificity of the employed PCR primers.

Genomic Features and Structure of the *O.vulgare* Mitogenome

The assembled genome data of *O.vulgare* is now publicly accessible in CNCB (China National Center for Bioinformation) (<https://www.cncb.ac.cn/>) [52]. Sequencing effort included approximately 12.9 Gb of PacBio HiFi long-read data and the sequencing yielded 432,565 clean reads. The resulting mitogenome assembly is a complex structure and a complete circular DNA structure spanning 497,243 bp was obtained after optimization (Figure 1C-D).

The mitogenome of *O.vulgare* contains a total of 57 annotated genes, including 29 core genes, seven plastid-derived genes, 14 transfer RNAs (tRNAs), and four ribosomal RNAs (rRNAs) (Figure 2A, Table 1). The PCGs are distributed across various functional categories: five genes related to ATP synthase (*atp1*, *atp4*, *atp6*, *atp8*, *atp9*), nine genes associated with NADH dehydrogenase (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *nad9*), and two gene encoding ubiquinol cytochrome c reductase (*cob*). Additionally, there are four cytochrome c biogenesis genes (*ccmB*, *ccmC*, *ccmFC*, *ccmFN*), three cytochrome c oxidase genes (*cox1*, *cox2*, *cox3*), a single maturase gene (*matR*). Other genes include succinate dehydrogenase membrane anchor subunit SDH4 (*sdh4*), three large subunit ribosomal protein genes (*rpl2*, *rpl5*, *rpl23*), and four small

subunit ribosomal protein genes (*rps4*, *rps12* (x2), *rps13*).

Codon Usage Analysis

The analysis of codon usage patterns in *O. vulgare* mitochondria revealed significant variation in synonymous codon preference, reflecting evolutionary adaptation for optimized translation efficiency. The observed bias corresponds with the tRNA pool availability, where preferred codons matching abundant tRNA species enable more efficient protein synthesis through facilitated ribosome progression. Examination of the 61 sense codons encoding 20 standard amino acids in mitochondrial PCGs (Figure 2B, Table S2) identified distinct codon preference patterns: codons AGA (Arg), GCU (Ala), CUU (Leu), UCU (Ser), ACU (Thr), GGA (Gly) and CCU (Pro) exhibited strong preferential usage (RSCU > 1.3), whereas CCG (Pro), CGU (Arg), CUG (Leu), GCG (Ala), and CGC (Arg) showed markedly reduced frequency (RSCU < 0.7). Notably, the codon repertoire demonstrated a moderate bias toward G/C nucleotides in the third wobble position, while termination codons displayed no significant usage preference. These findings suggest that translational selection has shaped the codon usage patterns in *O. vulgare* mitochondria to maximize the efficiency of the organellar translation machinery.

Repeat Elements Analysis

The mitogenome analysis revealed a diverse repertoire of repetitive elements, including nine simple sequence repeats (SSRs) with lengths ranging from 10 to 143 bp (Figure 3A, Table S3). These SSRs comprised three mononucleotide, three dinucleotide, two trinucleotide, and one compound repeat motifs. Additionally, we detected 26 tandem repeats (Table S4) and 607 dispersed repeats, consisting of 356 forward, 250 palindromic, and one reverse repeat units (Figure 3B, Table S5), with sizes varying from 30 to 48,238 bp. Particularly noteworthy were 14 large repeat sequences (>1 kb) spanning a total of 91,031 base pairs, which included nine palindromic (32,232 bp total) and five forward (58,799 bp total) repeats. These extensive repetitive elements are of special interest due to their potential role in mediating genomic rearrangements through

recombination events [53]. The spatial organization of these repetitive elements is illustrated in Figure 3C, with a circular representation showing: (1) SSRs in the outermost ring, (2) tandem repeats in the intermediate ring, and (3) dispersed repeats in the innermost connections, color-coded to distinguish forward (blue), palindromic (orange), and reverse (red) orientations.

Verification of Homologous Recombination

By examining the alignment of HiFi reads to potential recombinant structures, seven repeat pairs (r287, r227, r652, r510, P652, P1553 and P144) were detected as likely sites of homologous recombination. To verify the occurrence of these recombination events, primers were synthesized based on 100 bp regions adjacent to each repeat. PCR amplification using primer pairs (F1 & R1 and F2 & R2) successfully detected both the original genomic sequence and recombinant variants, as evidenced by the products generated from F1 & R2 and F2 & R1 primer combinations. The results demonstrate distinct electrophoretic bands, confirming structural variations resulting from recombination (Table 2, Figure 4A–G). Additionally, our previous research schematically depicts the recombination process mediated by forward repeats, highlighting how this mechanism contributes to mitochondrial DNA diversity [54].

Mitochondrial-Plastid DNA Transfer and Genome Characterization

The complete cpgenome of *O. vulgare* was successfully assembled into a circular DNA molecule of 151,833 bp (Figure 5A), exhibiting the typical quadripartite structure of angiosperm plastomes consisting of a large single-copy (LSC) region (83,046 bp), a small single-copy (SSC) region (17,597 bp), and two inverted repeat (IR) regions (25,595 bp each) (Figure 6A). Genome annotation identified a conserved set of photosynthetic genes including photosystem II components (psbA-C), photosystem I core subunits (psaA-B), the key Calvin cycle enzyme *rbcl*, and electron transport proteins (*petB*, *petD*) [55-59], along with essential genetic machinery genes such as RNA polymerase subunits (*rpoA*, *rpoB*, *rpoC1/C2*) that are crucial for chloroplast function [60, 61]. The structural organization and gene content of the *O. vulgare*

cpgenome conform to the conserved patterns observed in Lamiaceae species while providing species-specific genomic characteristics for comparative analyses (Figure 6B).

Comparative analysis identified nine mitochondrial plastid DNA sequences (MTPTs) resulting from potential intercompartmental gene transfer events between the chloroplast and mitogenomes (Table S6, Figure 5B). These transferred sequences, designated MTPT1 through MTPT109, exhibited considerable size variation ranging from 35 bp to 5,103 bp. The cumulative length of all MTPTs reached 71,010 bp, representing significant proportions of both organellar genomes - accounting for 14.28% of the mitogenome's total length and 46.77% of the cpgenome's sequence content. These findings demonstrate substantial genetic material exchange between the two organellar genomes in *O. vulgare*. Notably, 18 MTPTs contained complete tRNA genes: MTPT1 (*trnF-GAA* and *trnL-UAA*), MTPT4 (*trnL-CAA*), MTPT6 (*trnG-GCC* and *trnS-UGA*), MTPT19 (*trnI-CAU*), MTPT23 (*trnR-UCU*), MTPT27 (*trnG-GCC* in cpgenome), MTPT28 (*trnR-UCU* in cpgenome), MTPT31 (*trnI-CAU*), MTPT35 (*trnW-CCA* in cpgenome), MTPT38 (*trnP-UGG* in cpgenome), MTPT54 (*trnL-UAA* and *trnF-GAA* in cpgenome), MTPT60 (*trnM-CAU* in cpgenome), MTPT71 (*trnM-CAU* in cpgenome), MTPT78 (*trnL-CAA*), MTPT87 (*trnI-CAU*), MTPT100 (*trnP-UGG*), MTPT101 (*trnW-CCA*) and MTPT107 (*trnM-CAU*). Meanwhile, nine MTPTs harbored partial plastid-derived PCGs: MTPT3 (*psaB*), MTPT6 (*psbZ*), MTPT7 (*psbD*), MTPT10 (*rpoB*), MTPT40 (*ndhB*), MTPT90 (*infA*), MTPT100 (*psaJ*), MTPT102 (*petG*, *petL*, *psbE*, *psbF*, *psbL*) and MTPT109 (*ndhK* and *ndhJ*). These findings demonstrate distinct evolutionary fates for transferred genetic material: while tRNA sequences were maintained as complete, potentially functional units, PCGs exist only as fragmented remnants, suggesting their loss of biological activity through degenerative evolutionary processes following ancient DNA transfer events.

Variation and Prediction Analysis of RNA-editing sites

Comparative sequence analysis of eight conserved mitochondrial PCGs (*atp1*, *cox1*, *cytb*, *matR*, *nad1*, *nad4*, *rps12*, and *rps13*) across five angiosperm mitogenomes -

Lavandula angustifolia (NC_082307.1), *Platostoma chinense* (NC_069573.1), *O. vulgare* (this study), *Salvia rosmarinus* (PP992923.1), and *Salvia miltiorrhiza* (NC_023209.1) - revealed both striking evolutionary conservation and lineage-specific divergence patterns (Figure 7A-H). While alignment results demonstrated strong sequence conservation (>90% identity) in core functional regions, the identification of distinct nucleotide polymorphisms and amino acid substitutions (evident in non-green alignment regions) highlighted significant species-level variations. Notably, *O. vulgare* exhibited unexpected sequence divergence from its congener *S. rosmarinus* despite their shared Lamiales classification, with these genomic differences potentially reflecting either technical artifacts from divergent sequencing methodologies or authentic biological variation arising from distinct evolutionary trajectories and ecological adaptations. These findings not only provide quantitative evidence for mitogenome plasticity within Lamiaceae but also establish a framework for investigating structure-function relationships of mitochondrial genes, while highlighting important questions about the functional significance of these variations that warrant further study in ecological and evolutionary contexts.

The systematic investigation identified 395 RNA-editing sites across mitochondrial PCGs, all involving C-to-U conversions (Table S7-8; Figure 8A-B). The *nad4* gene displayed the most extensive RNA editing (36 sites), followed by *ccmB* (34 sites) and *ccmC* (30 sites). In contrast, several genes showed minimal editing: *nad6* (7 sites), *atp8* (6 sites), *atp1* (5 sites), *rps12* (4 sites), *rpl5* (4 sites), *rps3* (3 sites), and *rps13* (2 sites). Notably, all editing events occurred specifically at first or second codon positions, predominantly altering amino acid properties from hydrophilic to hydrophobic. The most frequent amino acid conversion was Ser to Leu (93 instances), followed by Pro to Leu (83 instances) and Ser to Phe (54 instances), while Thr to Met modifications were rare (2 occurrences). These findings demonstrate both the prevalence and functional implications of RNA editing in shaping mitochondrial protein diversity.

Evolutionary Relationships Among Closely Related Species

The maximum likelihood phylogenetic analysis, incorporating mitochondrial PCGs

from 25 representative species (using *Panax quinquefolius* [NC_067574.1] as the outgroup), robustly resolved the evolutionary relationships within Lamiaceae (Figure 8C). The resulting topology demonstrated strong nodal support (bootstrap >90%) for a sister-taxon relationship between *O. vulgare* and *Salvia rosmarinus* [PP992923.1], consistent with the established Angiosperm Phylogeny Group (APG) classification framework [62]. Comparative genomic examination further revealed dynamic patterns of mitochondrial gene evolution, showing complete retention of core respiratory genes across all taxa and lineage-specific losses of variable genes (particularly rps and rpl ribosomal protein genes) in certain clades (Figure 8D). These findings not only corroborate current phylogenetic hypotheses but also highlight the complex interplay between functional constraint and evolutionary plasticity in plant mitogenomes.

Selection variation of mitochondrial PCGs

Analysis for incorporating evolutionary selection pressure assessment through Ka/Ks ratio calculations for 18 mitochondrial PCGs in *O. vulgare* vs. 24 relatives, revealed significant variation in gene content (Figure 9A). Notably, several species exhibited incomplete PCG sets, with *Salvia rosmarinus* missing 7 PCGs, followed by *Lavandula angustifolia* and *Salvia miltiorrhiza* (each lacking 6 PCGs), along with *Platycodon grandiflorus* and *Platostoma chinense* showing partial gene loss (Table S9; Figure 5E). Selection pressure analysis of 18 PCGs in *Salvia rosmarinus* demonstrated that while the *matR* gene showed Ka/Ks > 1 (indicating positive selection), all remaining genes exhibited Ka/Ks < 1, revealing predominant negative selection that maintains evolutionary conservation through purifying selection pressures. Additionally, *cox3* and *nad3* in *Salvia miltiorrhiza*, *cox3* in *Platostoma chinense* and *Lavandula angustifolia*, *nad3* in *Salvia rosmarinus* had Ka/Ks=0. In contrast, *atp4* had Ka/Ks>1 in many speices (*Daucus carota subsp sativus*, *Scyphiphora hydrophyllacea*, *Helianthus occidentalis*, *Panax quinquefolius*, *Dystaenia ibukiensis*, *Chamaemelum nobile*, *Solanum muricatum*, *Hancornia speciosa*, *Trachelospermum jasminoides*, *Matricaria chamomilla*, *Solanum lycopersicum*, *Ipomoea aquatic*, *Mandragora caulescens*, *Eupatorium lindleyanum*, *Anethum foeniculum* and *Lycium ruthenicum*).

These results revealed differential selection pressures acting on mitochondrial PCGs across the 24 examined plant species. While most genes showed strong evolutionary constraint ($Ka/Ks \ll 1$), indicating purifying selection that preserves essential functions, notable exceptions were identified with $Ka/Ks > 1$, suggesting positive selection driving adaptive evolution in specific lineages. These contrasting patterns likely reflect that stringent functional requirements for core metabolic processes maintained by purifying selection, versus and adaptive modifications in particular genes that may confer lineage-specific physiological advantages. The presence of both evolutionary patterns within mitogenomes highlights their dual role in maintaining conserved cellular functions while permitting adaptive innovation in response to environmental pressures.

Nucleotide diversity (Pi) analysis of 30 mitochondrial PCGs across Lamiales, Solanales, and Gentianales revealed distinct evolutionary patterns, with Pi values ranging from 0 to 0.09878 (Figure 9B). The highest diversity was observed in Lamiales (*nad2*, $Pi=0.09878$) and Gentianales (*atp1*, $Pi=0.07033$), while Solanales exhibited consistently low diversity (0.00101-0.02878). Notably, Gentianales maintained $Pi>0.01$ for all genes except *cox1*, *nad1* and *rps13*, compared to 17 genes in Solanales and only 10 genes in Lamiales. These patterns likely reflect varying evolutionary pressures, with the particularly constrained diversity in Solanales and Lamiales suggesting either strong purifying selection, historical population bottlenecks, or recent evolutionary origins that have limited mitochondrial sequence variation in these lineages. The differential diversity patterns across orders highlight the complex interplay of evolutionary forces shaping mitogenome evolution in flowering plants.

Secondary Structures of Mitochondrial tRNAs and Synteny analysis

The mitochondrial tRNA repertoire of *O.vulgare* maintained complete structural integrity, with all tRNAs folding into characteristic cloverleaf secondary conformations (Figure 10). Particularly the genomic duplication of specific tRNA genes, where *trnQ-UUG* occurred as two native copies and *trnM-CAU* existed in three native copies. This redundancy pattern suggests either: (1) evolutionary selection for enhanced

translational capacity of these particular tRNAs, or (2) non-functional maintenance through inefficient deletion of duplicated sequences. The preservation of perfect cloverleaf structures across all tRNAs, including the duplicated copies, indicates strict structural constraints essential for their biological function in mitochondrial protein synthesis.

Comparative genomic analysis revealed both conserved and rearranged segments when examining syntenic blocks between *O. vulgare* and four Lamiales species (*Lavandula angustifolia*, *Platostoma chinense*, *Salvia rosmarinus*, and *Salvia miltiorrhiza*) (Table S10, Figure 11A). The synteny plots, distinguishing co-directional (grey) from inverted (green) regions, demonstrated limited conservation of extended collinear blocks, with most homologous regions being remarkably short (Figure 11B, Figure S5). Several evolutionary mechanisms may account for this fragmented synteny pattern. It might due to the frequent genomic rearrangements, lineage-specific evolutionary dynamics, post-polyploidization effects and rearrangement hotspots. These findings collectively suggest that mitogenome architecture in Lamiales has undergone extensive reorganization, with limited preservation of ancestral synteny. The observed patterns reflect both the inherent plasticity of plant mitogenomes and lineage-specific evolutionary trajectories that have reshaped genomic organization. The prevalence of short, discordantly arranged homologous fragments underscores the dynamic nature of mitogenome evolution in this plant group.

DISCUSSION

This study successfully addresses the long-standing difficulties in plant mitogenome assembly through the application of HiFi long-read sequencing, yielding the first fully characterized mitogenome of *O. vulgare*. As the foundational reference mitogenome for the *Origanum* genus, this work provided a critical genomic resource for evolutionary studies, enabled accurate comparative analyses across Lamiaceae and established a benchmark for future investigations of adaptive evolution in medicinal plants.

The study implemented a dual-marker DNA barcoding system utilizing the universal *rbcL* gene and the highly variable *matK* region to ensure precise molecular

identification of *O. vulgare*. The amplification strategy employed both standard and custom-designed primer sets, which demonstrated high specificity and amplification efficiency as verified by PCR product size analysis. Taxonomic validation was achieved through comparison with authenticated voucher specimens obtained from the National Plant Specimen Resource Center. The established protocol represents a reliable and reproducible framework for plant species authentication, with particular value for resolving taxonomically challenging species complexes and important applications in biodiversity conservation and ecological research. [63, 64].

The mitogenome of *O. vulgare* adopts a circular conformation, characterized by an abundance of repetitive elements. Comparative analysis with related Lamiales species (*L. angustifolia*, *P. chinense*, *S. rosmarinus*, and *S. miltiorrhiza*) reveals conserved structural features, while repetitive sequence-mediated recombination generates conformation variations [65]. Our analysis identified nine polymorphic simple sequence repeats (SSRs) that represent potential molecular markers for species authentication, population genetic studies, comparative genomic analyses and genetic mapping applications [66-68]. Furthermore, we detected 26 tandem repeat loci and 607 dispersed repeat elements, which may contribute to genomic rearrangement processes, evolutionary adaptation, transcriptional regulation and phenotypic plasticity [69]. These repetitive elements additionally show tendency to generate mutant lines in functional genomic studies [70].

Plant mitogenomes exhibit remarkable genomic plasticity through the frequent acquisition of exogenous DNA via intracellular gene transfer. This process generates mosaic genome architectures containing sequences of diverse origin, including plastid-derived DNA, nuclear genome sequences and horizontally transferred genetic material. Such promiscuous DNA incorporation represents a key driver of structural complexity and evolutionary innovation in plant mitogenomes. Notably, MTPTs represent a conserved feature of angiosperm genomes, serving as molecular fossils that record historical evolutionary events and ongoing genome dynamics [71]. In *O. vulgare*, there are 109 MTPTs with a combined length of 71,010 bp. These MTPTs represent 14.28% of the total mitogenome sequence and exhibit significant homology with 46.77% of the

cpgenome. The GC composition of these transferred sequences aligns with patterns characteristic of neutral evolutionary processes in angiosperm mitogenomes, suggesting their gradual accumulation through non-selective mechanisms. [72-75]. MTPTs can be categorized into two distinct functional classes: PCGs and tRNAs. Comparative analysis reveals that while the most PCGs exist as fragmented and non-functional, the tRNA sequences maintain their structural integrity. Emerging genomic evidence indicates these MTPTs may play significant evolutionary roles, potentially contributing to enhancing genomic plasticity, increasing transcriptional efficiency and post-transcriptional RNA modification processes [76].

In vascular plants, RNA editing occurs predominantly in organellar transcripts, exhibiting the highest prevalence in mitochondrial mRNAs and a lower frequency in cpgenomes. This post-transcriptional modification is facilitated by pentatricopeptide repeat (PPR) proteins encoded in the nuclear genome, which bind to specific cis-regulatory elements in organellar RNAs [77], and the most prevalent editing event involves C-to-U conversions in mitochondrial transcripts [78-80], which critically influence RNA splicing, evolutionary adaptation, and developmental processes. Due to the central role of mitochondria in plant growth and development [81], mitochondrial RNA editing has become a major focus of research, with emerging applications in precision plant breeding and genetic improvement [82]. In *O. vulgare*, 395 C-to-U RNA editing sites were revealed within mitochondrial PCGs. Strikingly, nearly half (48.35%) of these modifications resulted in hydrophilic-to-hydrophobic amino acid substitutions, whereas only 7.34% showed the opposite conversion pattern (Table S8). This substantial bias toward increased protein hydrophobicity aligns with conserved evolutionary trends observed in plant mitogenomes, suggesting functional conservation of this post-transcriptional modification mechanism [83-85].

In general, non-synonymous nucleotide substitutions introduce amino acid alterations that may impact protein structure and functional properties, thereby serving as potential substrates for adaptive evolution. Such modifications can confer either selective advantages or deleterious effects, depending on their phenotypic consequences. In contrast, synonymous substitutions maintain the original amino acid

sequence and are generally considered evolutionarily neutral, reflecting the underlying mutation rate. The ratio of non-synonymous to synonymous substitution rates (K_a/K_s) provides a robust metric for evaluating selection pressures, where values >1 , $=1$, and <1 indicate positive selection, neutral evolution, and purifying selection, respectively [86]. Selection pressure analysis (K_a/K_s) across 18 PCGs in 24 relatives revealed distinct evolutionary patterns. The majority of genes exhibited K_a/K_s ratios <1 , indicating predominant purifying selection and functional conservation. However, nine genes (*atp4*, *atp8*, *ccmB*, *ccmFC*, *ccmFN*, *cox3*, *matR*, *nad9*, and *rps13*) showed signatures of positive selection ($K_a/K_s > 1$), potentially reflecting adaptive responses to environmental challenges. Notably, genes associated with key mitochondrial functions, including ATP synthase complex assembly, cytochrome c maturation, NADH dehydrogenase activity, and ribosomal small subunit formation, demonstrating stronger positive selection signals in *O. vulgare* relative to related species. These findings suggest that while synonymous substitutions accumulate neutrally in plant genomes, essential functional genes remain constrained by purifying selection, maintaining critical cellular processes despite genomic variation [87]. Comparative evolutionary analysis using K_a/K_s ratio quantification provides a powerful approach for identifying functional divergence among orthologous PCGs across taxonomic groups. This method enables discrimination between distinct selection regimes, revealing strong purifying selection ($K_a/K_s \ll 1$) which maintains essential functions, neutral evolution ($K_a/K_s \approx 1$) which reflects unconstrained variation and positive selection ($K_a/K_s \gg 1$) which indicates adaptive evolution. Such analyses yield critical insights into molecular evolution patterns while facilitating the construction of more robust phylogenetic frameworks through detection of lineage-specific selective pressures.

Phylogenetic analysis of the complete mitogenomes of *O. vulgare* and related Lamiales species reveals significant evolutionary insights through resolution of phylogenetic relationships, quantification of genetic divergence and identification of lineage-specific adaptations. While clustering firmly within the Lamiales order, *O. vulgare* exhibits distinctive genomic signatures indicative of evolutionary specialization. These unique mitogenomic features likely resulted from the combined

effects of prolonged geographic isolation, ecological niche adaptation and population genetic processes, collectively driving its phylogenetic differentiation from congeneric species [88, 89]. These findings, when integrated with the distinctive cpgenome architecture of *O. vulgare*, provide novel insights into speciation mechanisms and organellar genome evolution within Lamiaceae. Our comparative mitogenomic approach achieves dual objectives. It elucidated the genomic basis of *O. vulgare*'s phylogenetic uniqueness and uncovered fundamental principles governing evolutionary radiation in Lamiales. The study demonstrates how complex interactions between genetic constraints and environmental pressures drive lineage diversification, thereby advancing our understanding of plant genome dynamics. Moreover, this work establishes a critical framework for future research directions, including deciphering molecular adaptation pathways, developing evolutionary-informed conservation strategies and harnessing genomic insights for crop improvement applications.

This study elucidates the genomic divergence patterns between *O. vulgare* and related species, highlighting the synergistic effects of evolutionary dynamics, ecological adaptation, and habitat-specific selection pressures. Our findings provide critical insights into the mechanisms governing genetic diversification within the Lamiales order by precisely delineating interspecific relationships and adaptive signatures. The established framework not only advances understanding of evolutionary pathways and ecological specialization, but also enables targeted investigations into plant-environment interactions and adaptive radiation processes. These results have important implications for both fundamental evolutionary biology and applied research in species conservation and utilization.

Abbreviations

MTPTs: mitochondrial plastid DNAs; ccm: cytochrome c biogenesis genes; Leu: Leucine; PCGs: protein coding genes; NADH: nicotinamide adenine dinucleotide hydride; IR: inverted repeat; MSA: multiple sequence alignment; PCR: polymerase chain reaction; rRNA: ribosomal RNA; ATP: adenosine triphosphate; Gln: Glutamine; LSC: large single-copy; APG: Angiosperm Phylogeny Group; Cys: Cysteine; Pro: Proline; CTAB: cetyltrimethylammonium bromide; DEPC: diethyl pyrocarbonate; tRNAs: Transfer RNAs; GC: guanine - cytosine; Gly: Glycine; RSCU: relative synonymous codon usage; bp: base pairs; Phe: Phenylalanine; RNA: ribonucleic acid; Ile: Isoleucine; Arg: Arginine; N50: median read length of the longest contigs; Val: Valine; SSRs: simple sequence repeats; SSC: small single-copy; Ala: Alanine; Ser: Serine; Met: Methionine; Lys: Lysine; Thr: Threonine; Glu: Glutamate; mt: mitochondrial contig.

Authors' Contributions

M. Qin and X. Gong designed the study and supervised the project. M. Qin performed the mitochondrial genome assembly, annotation, and data analysis. S. Yu and M. Tang conducted the sequencing and assisted in genome assembly. I. Outssint contributed to the analysis of RNA editing sites and performed the phylogenetic analysis. All authors contributed to manuscript writing and approved the final version.

Funding

This work was supported by research grants from Zhenjiang City Social Development Projects (SH2024080, SH2024038), Zhenjiang Science and Technology Plan Project (SH2023078) and Jiangsu University Medical Education Collaborative Innovation Fund (JDYY2023009). These contributions were essential for the successful completion of this work.

Data Availability

The mitochondrial genome of *O. vulgare* has been registered under accession number C_AA110995.1, and the chloroplast genome is available under accession number C_AA108664.1. For prompt access to the raw sequences or preliminary annotation datasets, researchers may contact the corresponding author via official institutional email. All such requests will be evaluated based on the institution's material transfer and data sharing policies.

Conflict of Interest Statement

The authors declare no conflicts of interest related to the publication of this research.

Clinical Trial Number

Not applicable.

Ethics Approval and Consent to Participate

In this study, diploid *O. vulgare* samples were obtained from publicly available seed sources. The collection of plant materials and subsequent experimental procedures complied with all relevant institutional, national, and international guidelines and regulations.

Consent for Publication

Not applicable.

REFERENCES

1. Nurzyńska-Wierdak R, Walasek-Janusz M: **Chemical Composition, Biological Activity, and Potential Uses of Oregano (*Origanum vulgare* L.) and Oregano Essential Oil.** *Pharmaceuticals (Basel, Switzerland)* 2025, **18**(2).
2. Ozkan G, Baydar H, Erbas S: **The influence of harvest time on essential oil composition, phenolic constituents and antioxidant properties of Turkish oregano (*Origanum onites* L.).** *Journal of the science of food and agriculture* 2010, **90**(2):205-209.
3. Antoniadou M, Rozos G, Vaou N, Zaralis K, Ersanli C, Alexopoulos A, Dadamogia A, Varzakas T, Tzora A, Voidarou CC: **Comprehensive Bio-Screening of Phytochemistry and Biological Capacity of Oregano (*Origanum vulgare*) and *Salvia triloba* Extracts against Oral Cariogenic and Food-Origin Pathogenic Bacteria.** *Biomolecules* 2024, **14**(6).
4. Al Bayati MHM, Cengiz MF, Kitis EC, Orcun %J *Journal of Essential Oil Research: Comparison of antimicrobial activities of oregano, lavender, sage, anise and clove extracts obtained by supercritical fluid carbon dioxide extraction and essential oils obtained by hydrodistillation.* 2024, **36**(3):258-270.
5. Bhagwat S, Haytowitz D, Holden J: **Oxygen Radical Absorbance Capacity (ORAC) of selected foods.** *US Department of Agriculture* 2010.
6. Ortiz MI, Cariño-Cortés R, Fernández-Martínez E, Muñoz-Pérez VM, Castañeda-Hernández G, González-García MP: **Peripheral Antinociception Induced by Carvacrol in the Formalin Test Involves the Opioid Receptor-NO-cGMP-K(+) Channel Pathway.** *Metabolites* 2025, **15**(5).
7. Arunasree KM: **Anti-proliferative effects of carvacrol on a human metastatic breast cancer cell line, MDA-MB 231.** *Phytomedicine : international journal of phytotherapy and phytopharmacology* 2010, **17**(8-9):581-588.
8. Møller IM, Rasmusson AG, Van Aken O: **Plant mitochondria - past, present and future.** *The Plant journal : for cell and molecular biology* 2021, **108**(4):912-959.
9. Hammani K, Giegé P: **RNA metabolism in plant mitochondria.** *Trends in plant science* 2014, **19**(6):380-389.
10. Arimura SI, Finkemeier I, Kühn K, Takenaka M: **Multilayered Regulation of Plastids and Mitochondria.** *Plant & cell physiology* 2024, **65**(4):473-476.
11. Horvath SE, Daum G: **Lipids of mitochondria.** *Progress in lipid research* 2013, **52**(4):590-614.
12. Van Aken O: **Mitochondrial redox systems as central hubs in plant metabolism and signaling.** *Plant physiology* 2021, **186**(1):36-52.
13. Liberatore KL, Dukowic-Schulze S, Miller ME, Chen C, Kianian SF: **The role of mitochondria in plant development and stress tolerance.** *Free radical biology & medicine* 2016, **100**:238-256.
14. Wang J, Kan S, Liao X, Zhou J, Tembrock LR, Daniell H, Jin S, Wu Z: **Plant organellar genomes: much done, much more to do.** *Trends in plant science* 2024, **29**(7):754-769.
15. Ali NA, Song W, Huang J, Wu D, Zhao X: **Recent advances and biotechnological applications of RNA metabolism in plant chloroplasts and mitochondria.** *Critical reviews in biotechnology* 2024, **44**(8):1552-1573.
16. Maier RM, Zeltz P, Kössel H, Bonnard G, Gualberto JM, Grienemberger JM: **RNA editing in plant mitochondria and chloroplasts.** *Plant molecular biology* 1996, **32**(1-2):343-365.
17. Liu J, Yan HF, Ge XJ: **The Use of DNA Barcoding on Recently Diverged Species in the Genus *Gentiana* (Gentianaceae) in China.** *PloS one* 2016, **11**(4):e0153008.

682 18. Ratnasingham S, Hebert PD: **bold: The Barcode of Life Data System**
683 (<http://www.barcodinglife.org>). *Molecular ecology notes* 2007, **7**(3):355-364.

684 19. Chen Y, Ye W, Zhang Y, Xu Y: **High speed BLASTN: an accelerated MegaBLAST search tool.**
685 *Nucleic acids research* 2015, **43**(16):7762-7768.

686 20. Bi C, Shen F, Han F, Qu Y, Hou J, Xu K, Xu LA, He W, Wu Z, Yin T: **PMAT: an efficient plant**
687 **mitogenome assembly toolkit using low-coverage HiFi sequencing data.** *Horticulture research*
688 2024, **11**(3):uhae023.

689 21. Wick RR, Schultz MB, Zobel J, Holt KE: **Bandage: interactive visualization of de novo genome**
690 **assemblies.** *Bioinformatics (Oxford, England)* 2015, **31**(20):3350-3352.

691 22. Li J, Ni Y, Lu Q, Chen H, Liu C: **PMGA: A plant mitochondrial genome annotator.** *Plant*
692 *communications* 2025, **6**(3):101191.

693 23. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S: **GeSeq -**
694 **versatile and accurate annotation of organelle genomes.** *Nucleic acids research* 2017,
695 **45**(W1):W6-w11.

696 24. Greiner S, Lehwark P, Bock R: **OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded**
697 **toolkit for the graphical visualization of organellar genomes.** *Nucleic acids research* 2019,
698 **47**(W1):W59-w64.

699 25. Beier S, Thiel T, Münch T, Scholz U, Mascher M: **MISA-web: a web server for microsatellite**
700 **prediction.** *Bioinformatics (Oxford, England)* 2017, **33**(16):2583-2585.

701 26. Benson G: **Tandem repeats finder: a program to analyze DNA sequences.** *Nucleic acids*
702 *research* 1999, **27**(2):573-580.

703 27. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R: **REPuter: the**
704 **manifold applications of repeat analysis on a genomic scale.** *Nucleic acids research* 2001,
705 **29**(22):4633-4642.

706 28. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA: **Circos:**
707 **an information aesthetic for comparative genomics.** *Genome research* 2009, **19**(9):1639-1645.

708 29. Hershberg R, Petrov DA: **Selection on codon bias.** *Annual review of genetics* 2008, **42**:287-299.

709 30. Gouy M, Gautier C: **Codon usage in bacteria: correlation with gene expressivity.** *Nucleic acids*
710 *research* 1982, **10**(22):7055-7074.

711 31. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT: **PhyloSuite: An integrated and**
712 **scalable desktop platform for streamlined molecular sequence data management and**
713 **evolutionary phylogenetics studies.** *Molecular ecology resources* 2020, **20**(1):348-355.

714 32. Kolmogorov M, Yuan J, Lin Y, Pevzner PA: **Assembly of long, error-prone reads using repeat**
715 **graphs.** *Nature biotechnology* 2019, **37**(5):540-546.

716 33. Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C: **CPGAVAS2, an integrated plastome**
717 **sequence annotator and analyzer.** *Nucleic acids research* 2019, **47**(W1):W65-w73.

718 34. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I: **VISTA: computational tools for**
719 **comparative genomics.** *Nucleic acids research* 2004, **32**(Web Server issue):W273-279.

720 35. Dubchak I, Brudno M, Loots GG, Pachter L, Mayor C, Rubin EM, Frazer KA: **Active conservation**
721 **of noncoding sequences revealed by three-way species comparisons.** *Genome research* 2000,
722 **10**(9):1304-1306.

723 36. Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer KA, Pachter LS, Dubchak I: **VISTA :**
724 **visualizing global DNA sequence alignments of arbitrary length.** *Bioinformatics (Oxford,*
725 *England)* 2000, **16**(11):1046-1047.

726 37. Brudno M, Malde S, Poliakov A, Do CB, Couronne O, Dubchak I, Batzoglu S: **Glocal alignment:
727 finding rearrangements during alignment.** *Bioinformatics (Oxford, England)* 2003, **19 Suppl**
728 **1:i54-62.**

729 38. Wang XC, Chen H, Yang D, Liu C: **Diversity of mitochondrial plastid DNAs (MTPTs) in seed
730 plants.** *Mitochondrial DNA Part A, DNA mapping, sequencing, and analysis* 2018, **29(4):635-**
731 **642.**

732 39. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R: **TBtools: An Integrative Toolkit
733 Developed for Interactive Analyses of Big Biological Data.** *Molecular plant* 2020, **13(8):1194-**
734 **1202.**

735 40. Lukeš J, Kaur B, Speijer D: **RNA Editing in Mitochondria and Plastids: Weird and Widespread.**
736 *Trends in genetics : TIG* 2021, **37(2):99-102.**

737 41. Lenz H, Hein A, Knoop V: **Plant organelle RNA editing and its specificity factors: enhancements
738 of analyses and new database features in PREPACT 3.0.** *BMC bioinformatics* 2018, **19(1):255.**

739 42. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ: **Jalview Version 2--a multiple
740 sequence alignment editor and analysis workbench.** *Bioinformatics (Oxford, England)* 2009,
741 **25(9):1189-1191.**

742 43. Chan PP, Lin BY, Mak AJ, Lowe TM: **tRNAscan-SE 2.0: improved detection and functional
743 classification of transfer RNA genes.** *Nucleic acids research* 2021, **49(16):9077-9096.**

744 44. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J: **KaKs_Calculator 2.0: a toolkit incorporating gamma-
745 series methods and sliding window strategies.** *Genomics, proteomics & bioinformatics* 2010,
746 **8(1):77-80.**

747 45. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE,
748 Sánchez-Gracia A: **DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets.**
749 *Molecular biology and evolution* 2017, **34(12):3299-3302.**

750 46. Katoh K, Standley DM: **MAFFT multiple sequence alignment software version 7:
751 improvements in performance and usability.** *Molecular biology and evolution* 2013,
752 **30(4):772-780.**

753 47. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R:
754 **IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era.**
755 *Molecular biology and evolution* 2020, **37(5):1530-1534.**

756 48. Letunic I, Bork P: **Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display
757 and annotation.** *Bioinformatics (Oxford, England)* 2007, **23(1):127-128.**

758 49. He W, Yang J, Jing Y, Xu L, Yu K, Fang X: **NGenomeSyn: an easy-to-use and flexible tool for
759 publication-ready visualization of syntenic relationships across multiple genomes.**
760 *Bioinformatics (Oxford, England)* 2023, **39(3).**

761 50. Darling AC, Mau B, Blattner FR, Perna NT: **Mauve: multiple alignment of conserved genomic
762 sequence with rearrangements.** *Genome research* 2004, **14(7):1394-1403.**

763 51. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL: **Versatile and
764 open software for comparing large genomes.** *Genome biology* 2004, **5(2):R12.**

765 52. **Database Resources of the National Genomics Data Center, China National Center for
766 Bioinformation in 2024.** *Nucleic acids research* 2024, **52(D1):D18-d32.**

767 53. Bi C, Paterson AH, Wang X, Xu Y, Wu D, Qu Y, Jiang A, Ye Q, Ye N: **Analysis of the Complete
768 Mitochondrial Genome Sequence of the Diploid Cotton *Gossypium raimondii* by
769 Comparative Genomics Approaches.** *BioMed research international* 2016, **2016:5040598.**

- 770 54. Wang Z, Wang H, Gong X, Ou X, Guo Y, Tang M: **De Novo assembly and phylogenetic analysis**
771 **of the complete mitochondrial genome of Eleutherococcus senticosus and related**
772 **araliaceous species.** *BMC plant biology* 2025, **25**(1):760.
- 773 55. Williams-Carrier R, Chotewutmontri P, Perkel S, Rojas M, Belcher S, Barkan A: **The psbA ORF**
774 **acts in cis to toggle HCF173 from an activator to a repressor for light-regulated psbA**
775 **translation in plants.** *The Plant cell* 2025.
- 776 56. Yan Z, Li X, Li Z, Song J, Dou H, Yang Y, Zhang G: **How to utilize far-red photons effectively:**
777 **substitution or supplementation with photosynthetically active radiation? A case study of**
778 **greenhouse lettuce.** *BMC plant biology* 2025, **25**(1):228.
- 779 57. Zhou X, Wang M, Yang L, Wang W, Zhang Y, Liu L, Chai J, Liu H, Zhao G: **Comparative**
780 **Physiological and Transcriptomic Analyses of Oat (Avena sativa) Seedlings under Salt Stress**
781 **Reveal Salt Tolerance Mechanisms.** *Plants (Basel, Switzerland)* 2024, **13**(16).
- 782 58. Vijayan S, Liu R, George S, Bhaskaran S: **Polyethylene terephthalate nanoparticles induce**
783 **oxidative damage in Chlorella vulgaris.** *Plant physiology and biochemistry : PPB* 2024,
784 **215**:108987.
- 785 59. Li N, Wong WS, Feng L, Wang C, Wong KS, Zhang N, Yang W, Jiang Y, Jiang L, He JX: **The thylakoid**
786 **membrane protein NTA1 is an assembly factor of the cytochrome b(6)f complex essential for**
787 **chloroplast development in Arabidopsis.** *Plant communications* 2023, **4**(1):100509.
- 788 60. Cho MS, Yang J, Kim SH, Crawford DJ, Stuessy TF, López-Sepúlveda P, Kim SC: **Plastid**
789 **phylogenomics of Robinsonia (Senecioneae; Asteraceae), endemic to the Juan Fernández**
790 **Islands: insights into structural organization and molecular evolution.** *BMC plant biology* 2024,
791 **24**(1):1016.
- 792 61. Xu X, Huang H, Lin S, Zhou L, Yi Y, Lin E, Feng L, Zheng Y, Lin A, Yu L *et al*: **Twelve newly**
793 **assembled jasmine chloroplast genomes: unveiling genomic diversity, phylogenetic**
794 **relationships and evolutionary patterns among Oleaceae and Jasminum species.** *BMC plant*
795 *biology* 2024, **24**(1):331.
- 796 62. Group TAP: **An update of the Angiosperm Phylogeny Group classification for the orders and**
797 **families of flowering plants: APG IV.** 2016, **181**(1):1-20.
- 798 63. Zhou G, Qin M, Liu X, Qi Y, Ou X, Tang M: **De novo assembly of the mitochondrial genome of**
799 **Glycyrrhiza glabra and identification of two types of homologous recombination**
800 **configurations caused by repeat sequences.** *BMC genomics* 2025, **26**(1):13.
- 801 64. Gong X, Zhang H, Guo Y, Yu S, Tang M: **Chromosome-level genome assembly of Iodes seguinii**
802 **and its metabonomic implications for rheumatoid arthritis treatment.** *The plant genome*
803 2025, **18**(1):e20534.
- 804 65. Ding H, Bi D, Zhang S, Han S, Ye Y, Yi R, Yang J, Liu B, Wu L, Zhuo R *et al*: **The Mitogenome of**
805 **Sedum plumbizincicola (Crassulaceae): Insights into RNA Editing, Lateral Gene Transfer, and**
806 **Phylogenetic Implications.** *Biology* 2022, **11**(11).
- 807 66. Qi WH, Jiang XM, Du LM, Xiao GS, Hu TZ, Yue BS, Quan QM: **Genome-Wide Survey and Analysis**
808 **of Microsatellite Sequences in Bovid Species.** *PloS one* 2015, **10**(7):e0133667.
- 809 67. Behura SK, Severson DW: **Genome-wide comparative analysis of simple sequence coding**
810 **repeats among 25 insect species.** *Gene* 2012, **504**(2):226-232.
- 811 68. Miao XX, Xub SJ, Li MH, Li MW, Huang JH, Dai FY, Marino SW, Mills DR, Zeng P, Mita K *et al*:
812 **Simple sequence repeat-based consensus linkage map of Bombyx mori.** *Proceedings of the*
813 *National Academy of Sciences of the United States of America* 2005, **102**(45):16303-16308.

69. Lisch D: **How important are transposons for plant evolution?** *Nature reviews Genetics* 2013, **14**(1):49-61.
70. Veilleux RE, Mills KP, Baxter AJ, Upham KT, Ferguson TJ, Holt SH, Lu N, Ruiz-Rojas JJ, Pantazis CJ, Davis CM *et al*: **Transposon tagging in diploid strawberry.** *Plant biotechnology journal* 2012, **10**(8):985-994.
71. Wei L, Liu TJ, Hao G, Ge XJ, Yan HF: **Comparative analyses of three complete *Primula* mitogenomes with insights into mitogenome size variation in Ericales.** *BMC genomics* 2022, **23**(1):770.
72. Sloan DB, Wu Z: **History of plastid DNA insertions reveals weak deletion and at mutation biases in angiosperm mitochondrial genomes.** *Genome biology and evolution* 2014, **6**(12):3210-3221.
73. Zhu H, Shan Y, Li J, Zhang X, Yu J, Wang H: **Assembly and comparative analysis of the complete mitochondrial genome of *Viburnum chinshanense*.** *BMC plant biology* 2023, **23**(1):487.
74. Sun C, Chen Y, Zheng D, Zhong Y, Luo S, Meng S, Qian L, Wei D, Liu Y, Dai S *et al*: **The complete mitochondrial genome of *Bauhinia variegata* (Leguminosae).** *Mitochondrial DNA Part B, Resources* 2024, **9**(1):128-132.
75. Wang D, Wu YW, Shih AC, Wu CS, Wang YN, Chaw SM: **Transfer of chloroplast genomic DNA to mitochondrial genome occurred at least 300 MYA.** *Molecular biology and evolution* 2007, **24**(9):2040-2048.
76. Wang D, Rousseau-Gueutin M, Timmis JN: **Plastid sequences contribute to some plant mitochondrial genes.** *Molecular biology and evolution* 2012, **29**(7):1707-1711.
77. Qiu F, Xue C, Liu J, Li B, Gao Q, Liang R, Chen K, Gao C: **An efficient mRNA delivery system for genome editing in plants.** *Plant biotechnology journal* 2025, **23**(4):1348-1358.
78. Sun Y, Xie M, Xu Z, Chan KC, Zhong JY, Fan K, Wong-Bajracharya J, Lam HM, Lim BL: **Differential RNA Editing and Intron Splicing in Soybean Mitochondria during Nodulation.** *International journal of molecular sciences* 2020, **21**(24).
79. Edera AA, Gandini CL, Sanchez-Puerta MV: **Towards a comprehensive picture of C-to-U RNA editing sites in angiosperm mitochondria.** *Plant molecular biology* 2018, **97**(3):215-231.
80. Giegé P, Brennicke A: **RNA editing in Arabidopsis mitochondria effects 441 C to U changes in ORFs.** *Proceedings of the National Academy of Sciences of the United States of America* 1999, **96**(26):15324-15329.
81. He P, Xiao G, Liu H, Zhang L, Zhao L, Tang M, Huang S, An Y, Yu J: **Two pivotal RNA editing sites in the mitochondrial *atp1*mRNA are required for ATP synthase to produce sufficient ATP for cotton fiber cell elongation.** *The New phytologist* 2018, **218**(1):167-182.
82. Yang J, Yang X, Su T, Hu Z, Zhang M: **The Development of Mitochondrial Gene Editing Tools and Their Possible Roles in Crop Improvement for Future Agriculture.** *Advanced genetics (Hoboken, NJ)* 2022, **3**(1):2100019.
83. Wu CS, Chaw SM: **Evolution of mitochondrial RNA editing in extant gymnosperms.** *The Plant journal : for cell and molecular biology* 2022, **111**(6):1676-1687.
84. Rehman O, Uzair M, Chao H, Khan MR, Chen M: **Decoding RNA Editing Sites Through Transcriptome Analysis in Rice Under Alkaline Stress.** *Frontiers in plant science* 2022, **13**:892729.
85. Kaila T, Saxena S, Ramakrishna G, Tyagi A, Tribhuvan KU, Srivastava H, Sandhya, Chaudhury A, Singh NK, Gaikwad K: **Comparative RNA editing profile of mitochondrial transcripts in**

858 **cytoplasmic male sterile and fertile pigeonpea reveal significant changes at the protein level.**
859 *Molecular biology reports* 2019, **46**(2):2067-2084.

860 86. Song Y, Du X, Li A, Fan A, He L, Sun Z, Niu Y, Qiao Y: **Assembly and analysis of the complete**
861 **mitochondrial genome of Forsythia suspensa (Thunb.) Vahl.** *BMC genomics* 2023, **24**(1):708.

862 87. Mower JP, Touzet P, Gummow JS, Delph LF, Palmer JD: **Extensive variation in synonymous**
863 **substitution rates in mitochondrial genes of seed plants.** *BMC evolutionary biology* 2007,
864 **7**:135.

865 88. Liu H, Yu J, Yu X, Zhang D, Chang H, Li W, Song H, Cui Z, Wang P, Luo Y *et al*: **Structural variation**
866 **of mitochondrial genomes sheds light on evolutionary history of soybeans.** *The Plant journal :*
867 *for cell and molecular biology* 2021, **108**(5):1456-1472.

868 89. Chen F, Song Y, Li X, Chen J, Mo L, Zhang X, Lin Z, Zhang L: **Genome sequences of horticultural**
869 **plants: past, present, and future.** *Horticulture research* 2019, **6**:112.

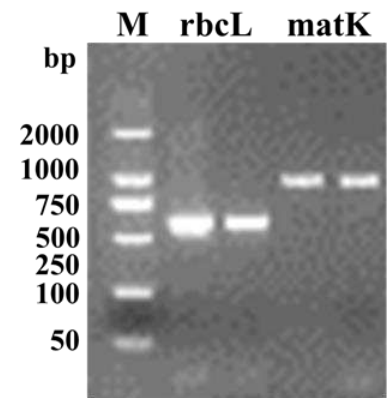
870

Figure 1 | Identification and assembly map of *O.vulgare*. (A) Morphological characteristics of *O.vulgare*. The left panel shows the specimen we collected, while the right panel is from the National Plant Specimen Resource Center of China Digital Herbarium, with the specimen number KUN1566203. The physical specimen is stored at the Herbarium, Kunming Institute of Botany, Chinese Academy of Sciences. (B) The *rbcL* and *matK* gene was amplified and compared by DNA barcoding technique. (C) Graphical Fragment Assembly (GFA) map of *O.vulgare* before unlooping. (D) GFA map of *O.vulgare* after unlooping.

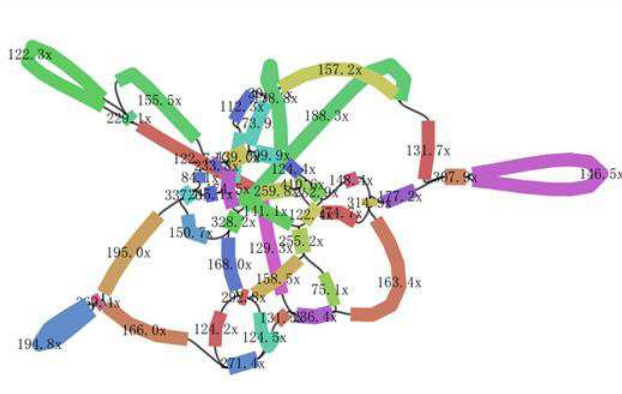
A



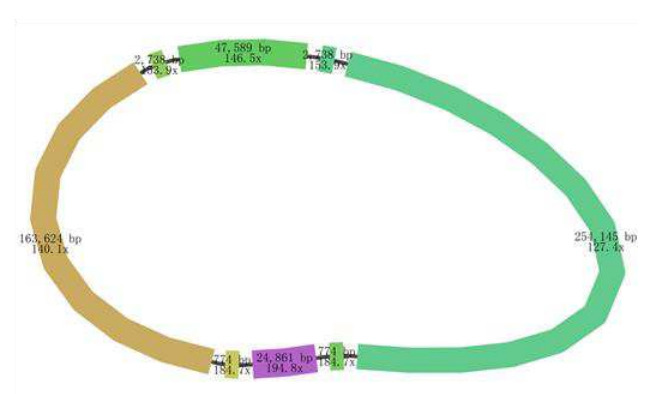
B



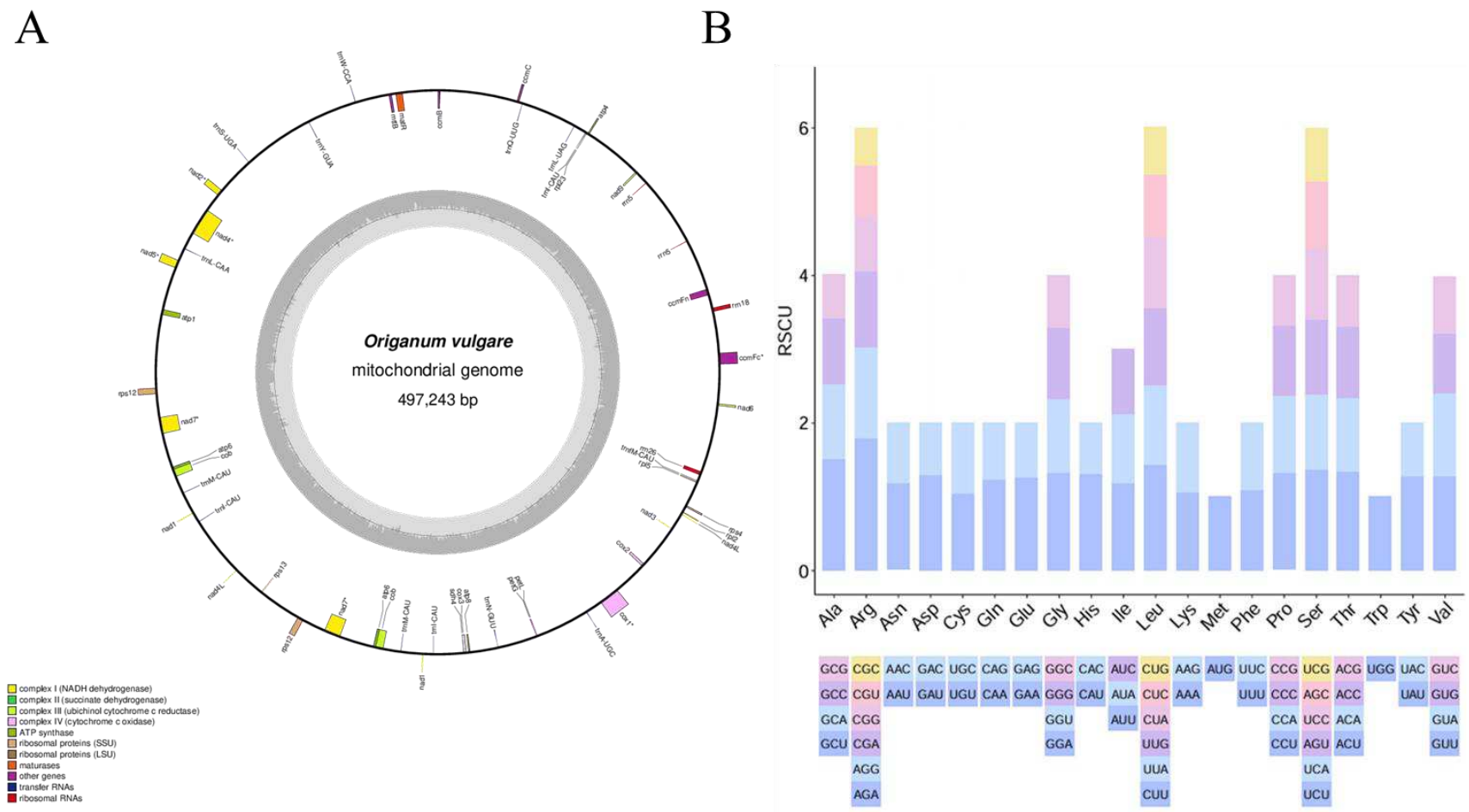
C



D



912 **Figure 2 | Mitochondrial annotation and Relative synonymous codon usage (RSCU) map of *O.vulgare*.** (A) Genes located on the outer side
913 of the circle are transcribed in a clockwise direction, whereas those on the inner side are transcribed counterclockwise. The dark grey plot on the
914 innermost circle represents the GC content. (B) Codon usage preference of *O.vulgare* mitogenome.



915 **Figure 3 | Repeat elements of the *O. vulgare* mitogenome.** (A) Simple sequence repeats (SSRs) (B) Dispersed repeats. (C) It shows the distribution
 916 of repeated elements, where the outermost circle represents SSRs, followed by tandem repeats, and the innermost connecting lines indicate
 917 dispersed repeats. The green lines depict forward repeats, the orange lines represent palindromic repeats, and the red lines illustrate reverse repeats.
 918

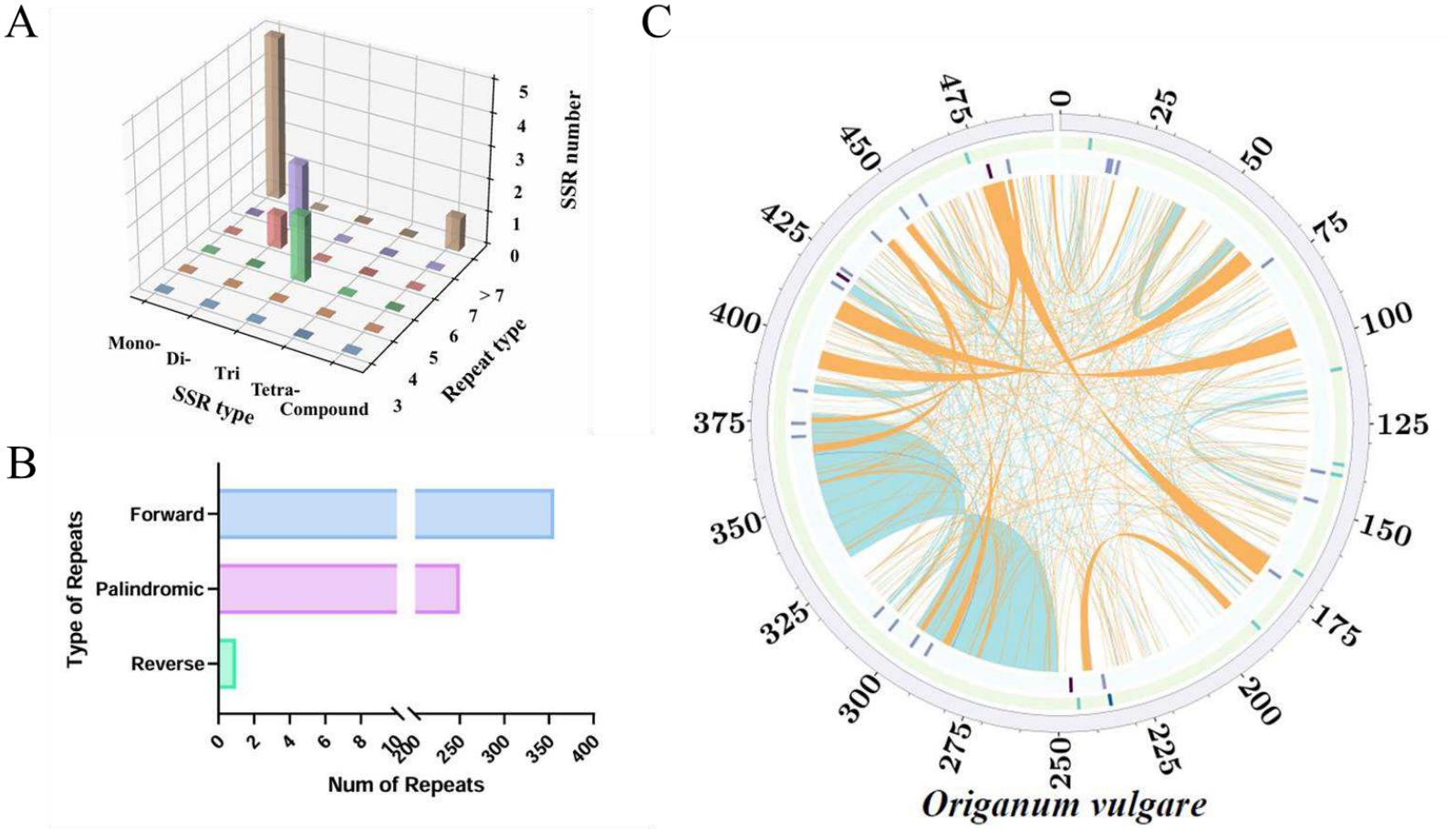
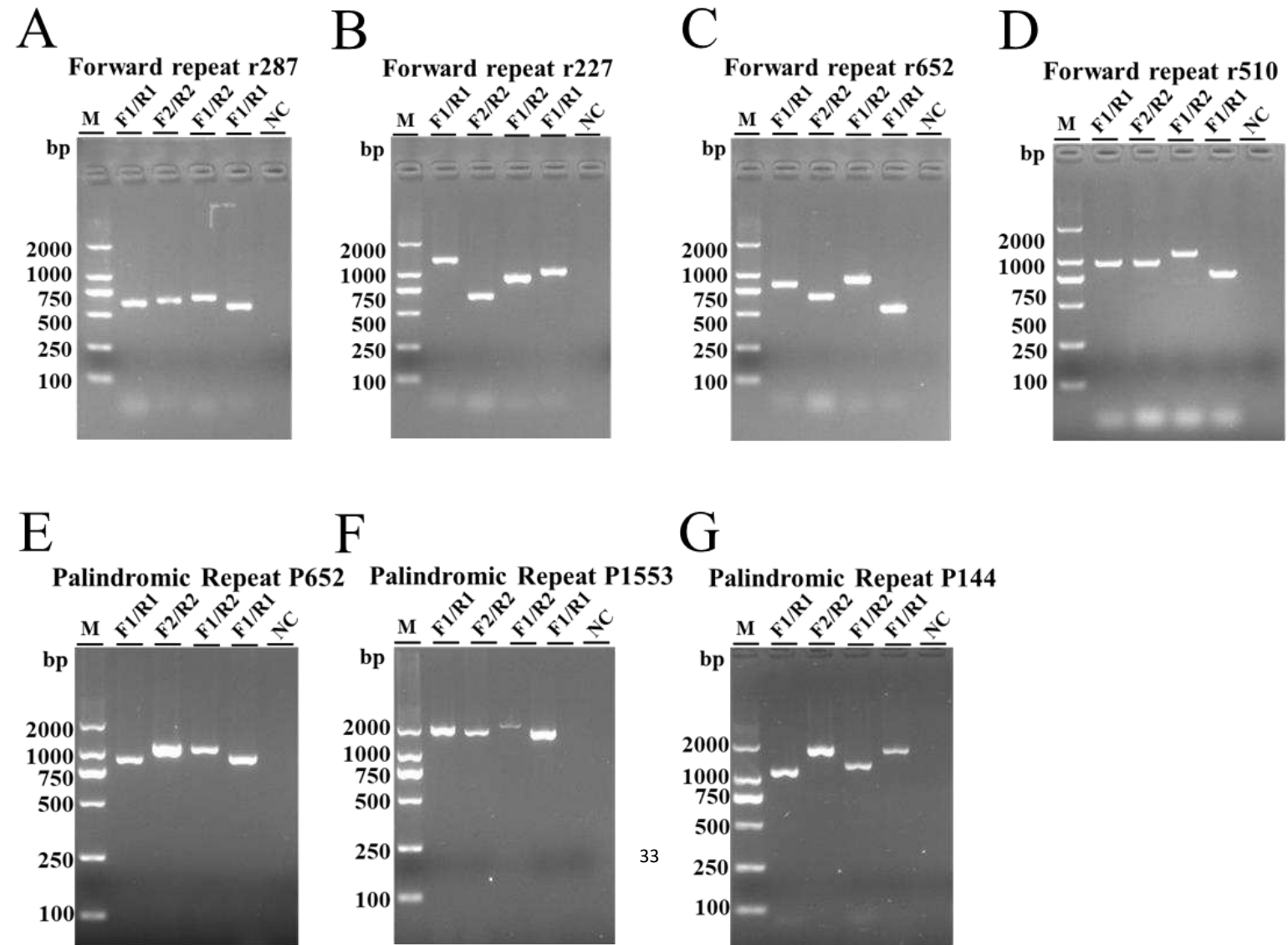


Figure 4 | Structural diversity of mitochondrial DNA driven by repeat-mediated recombination events. (A-G) PCR results showcase the multiple configurations of mitochondrial DNA, with each panel displaying six lanes: (1) molecular marker, (2) and (3) two major conformations, and (4) and (5) two minor conformations, (6) control, corresponding to recombination within seven repeat sequences (r287, r227, r652, r510, P652, P1553 and P144). The PCR bands clearly demonstrate the structural differences among these conformations.



945 **Figure 5 | Cpgenome Map and MTPTs of *O. vulgare*.** (A) Cpgenome map of *O. vulgare*. (B) MTPTs of *O. vulgare*, showcasing regions of sequence
 946 transfer between the chloroplast and mitogenomes. (C) The regional boundaries of chloroplasts of *O. vulgare* and other 3 relative species. (D) The
 947 Sequence identity plot of five Lamiales species.

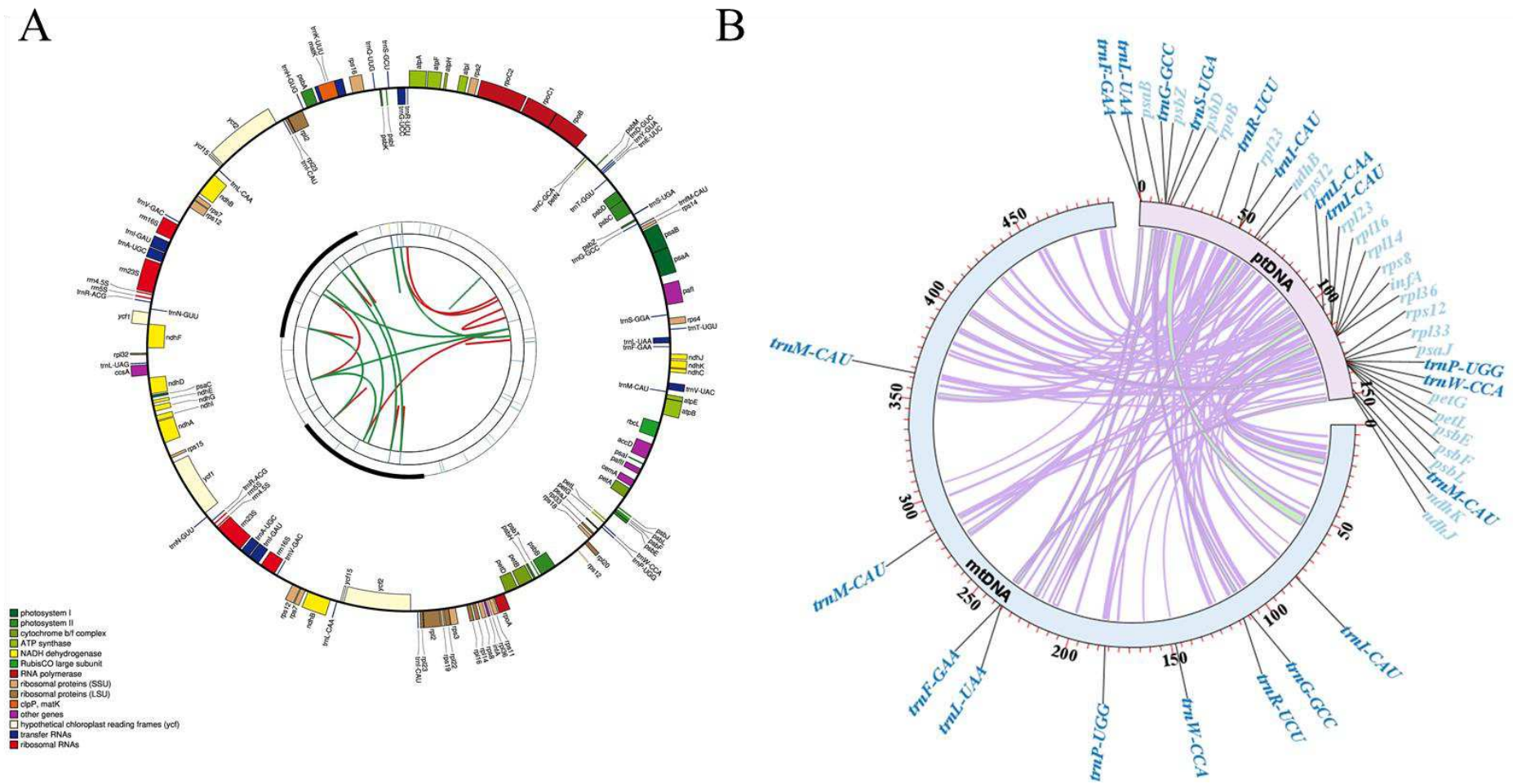


Figure 6 | Comparative analyses of cpgenome. (A) The regional boundaries of chloroplasts of *O. vulgare* and other 4 relative species. (B) The sequence identity plot of five Lamiales species.

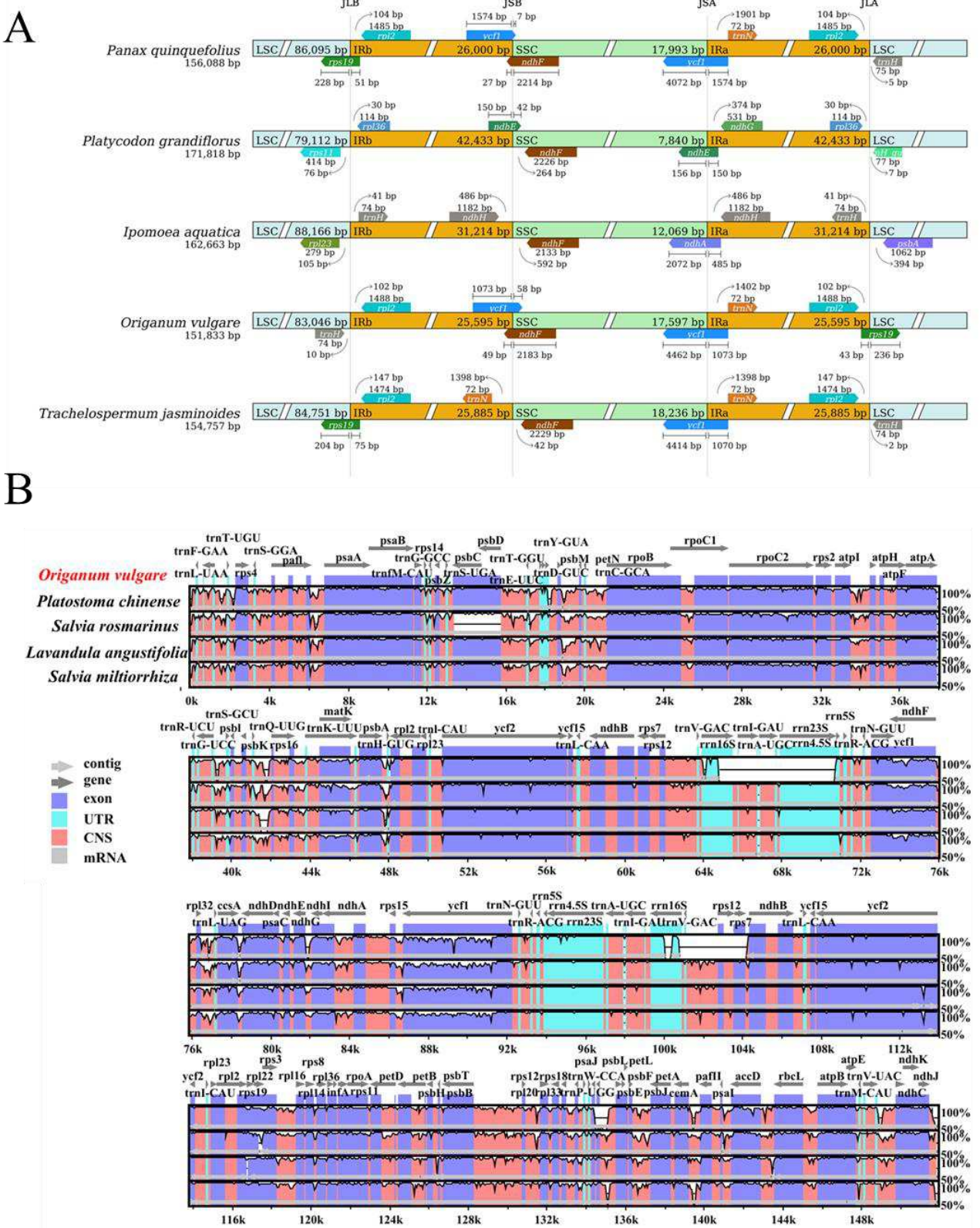


Figure 7 | Alignments of the nucleotide and amino acid sequences of PCGs from five mitogenomes. Multiple sequence alignments were conducted using MAFFT software with default parameters and visualized with ESPrpt 3.0. Panels (A-H) correspond to the alignment results of the *atp1*, *ccmB*, *cox1*, *matR*, *nad4*, *nad4L*, *nad7*, *rpl5* and *rps12* genes, respectively.

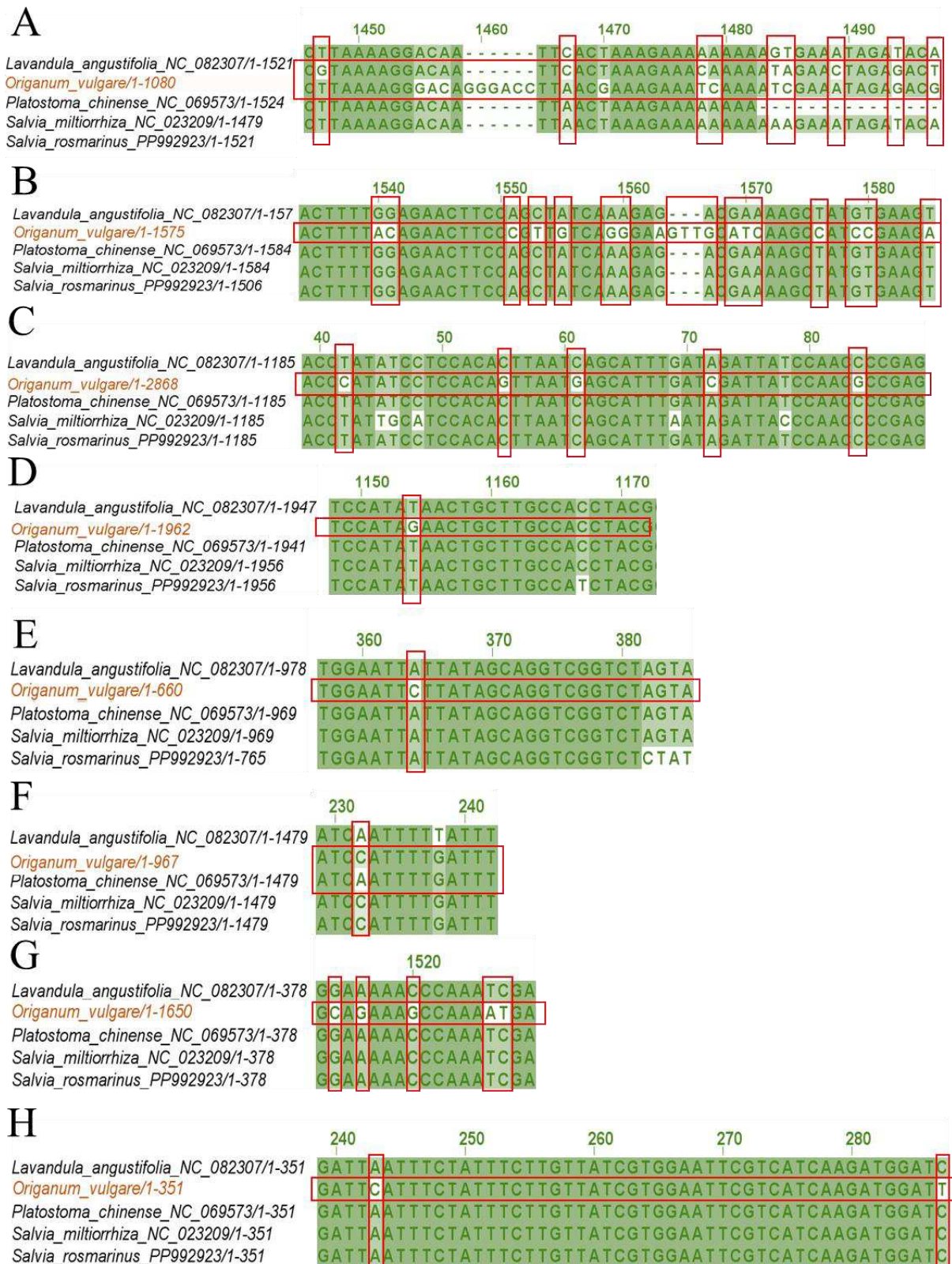


Figure 8 | RNA-editing sites and phylogenetic analysis based on mitochondrial PCGs from 25 species. (A) Statistics on the number of amino acid changes in the mitochondrial PCGs of *O.vulgare*. (B) Statistics on the number of RNA-editing sites in the mitochondrial PCGs of *O.vulgare*. (C) The number on the evolutionary tree indicates the credibility of a branch. (D) Distribution of PCGs in mitochondrial genomes of *O.vulgare* and its closely related species.

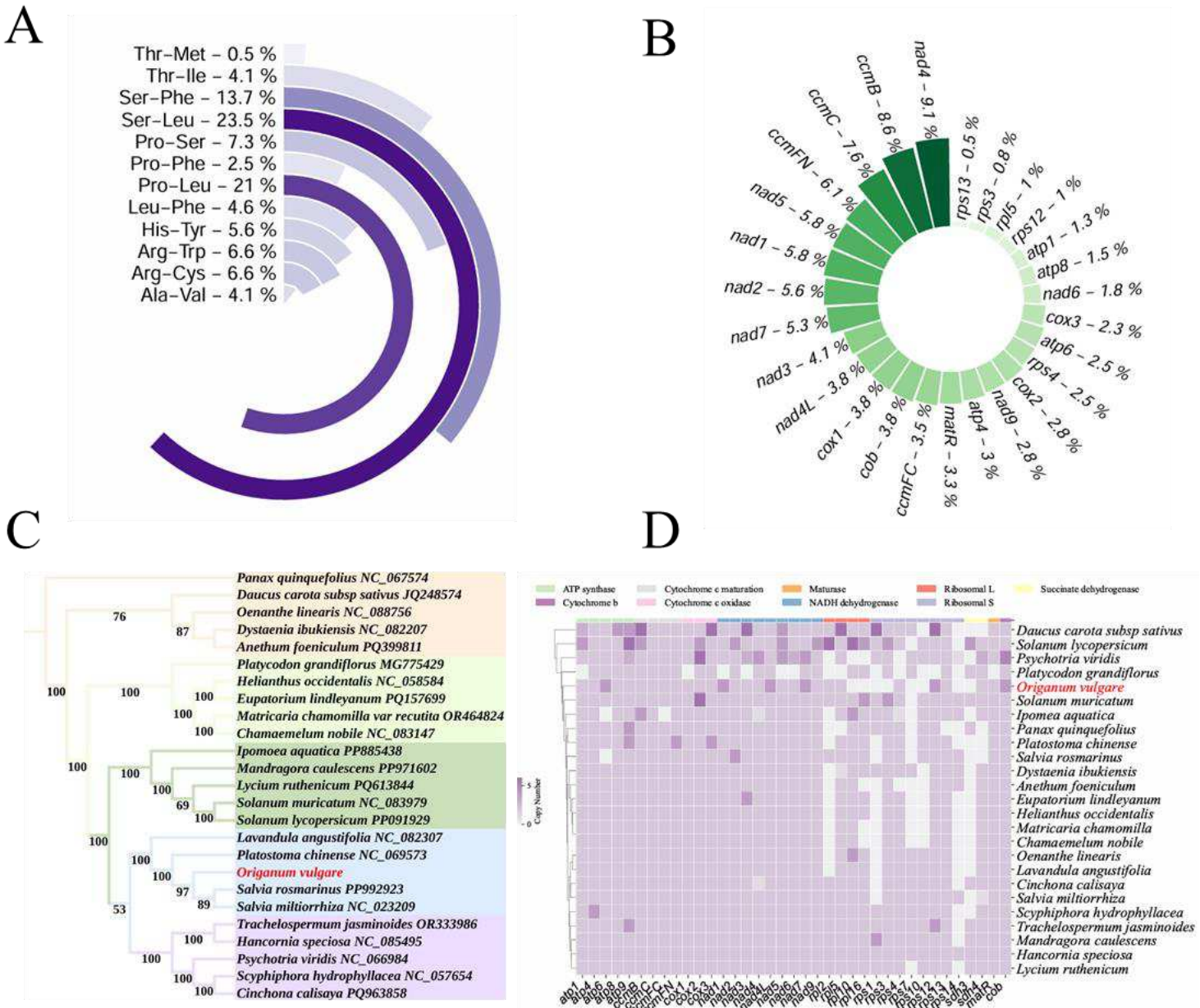


Figure 9 | KaKs and Nucleotide diversity (Pi) analysis of 25 speices. (A) Selective pressure analysis based on 18 PCGs. (B) Nucleotide diversity analysis of Lamiales, Solanales and Gentianales, each based on 5 species.

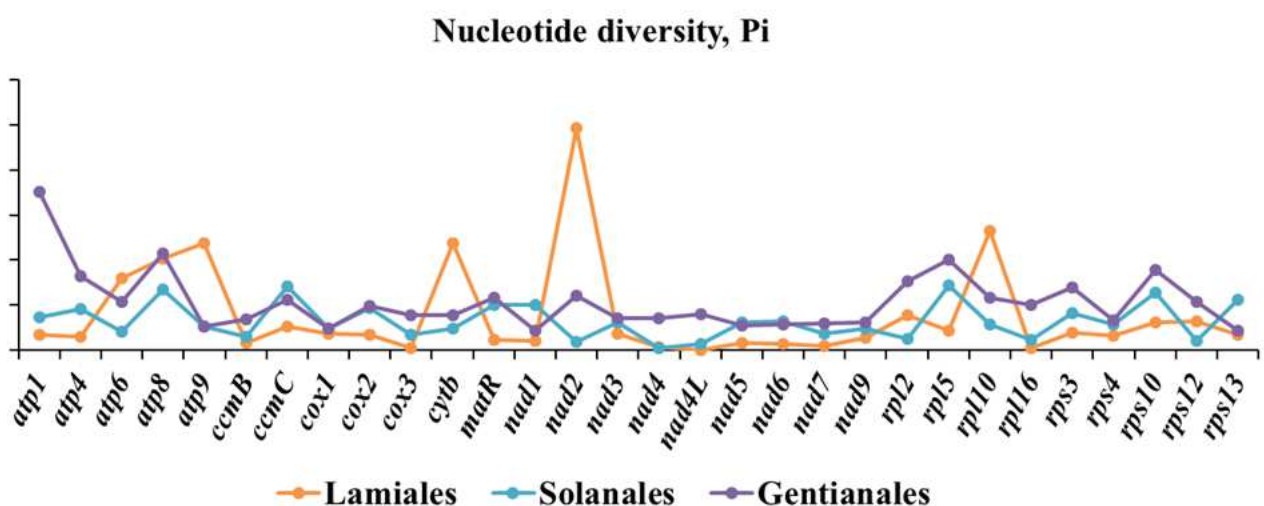
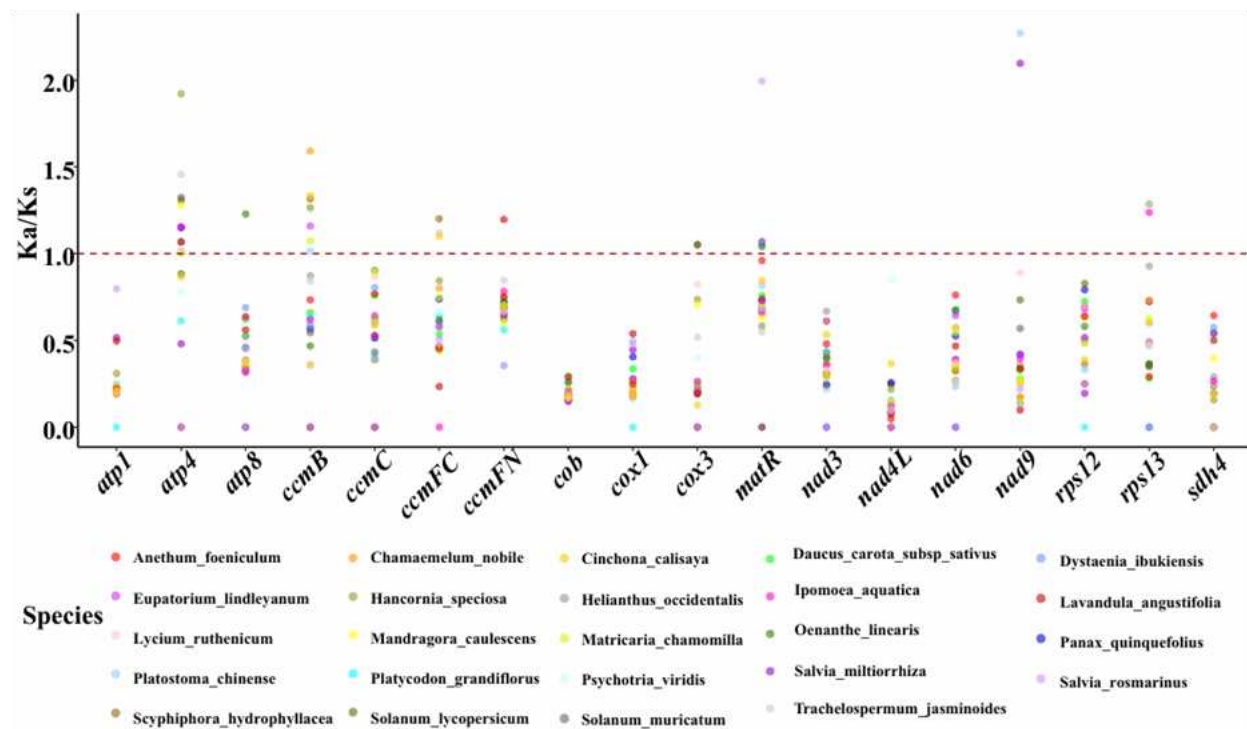


Figure 10 | The predicted secondary structures of tRNAs found in mitogenome of *O.vulgare*. Predicted secondary structures of tRNAs in mitogenome.

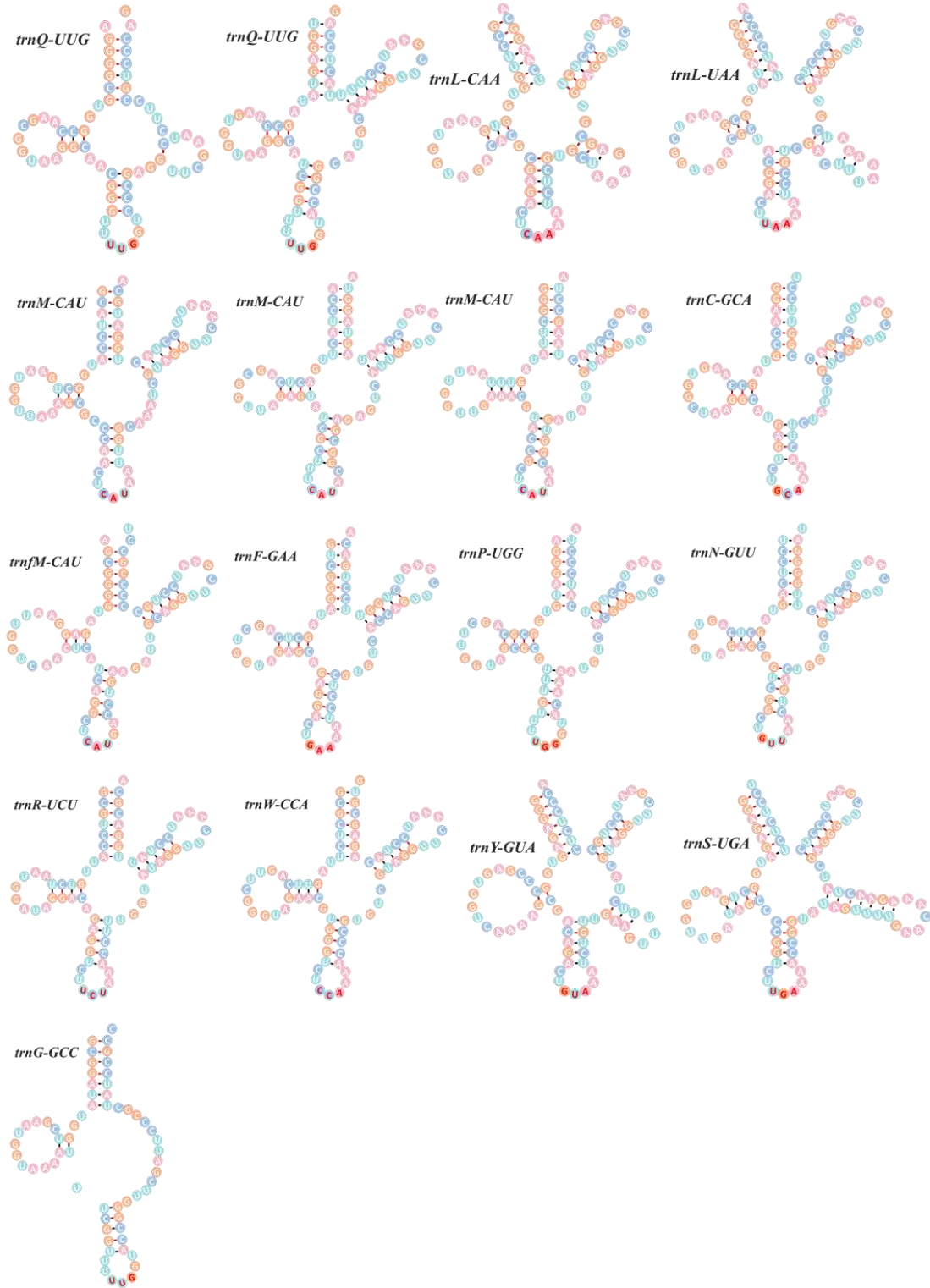
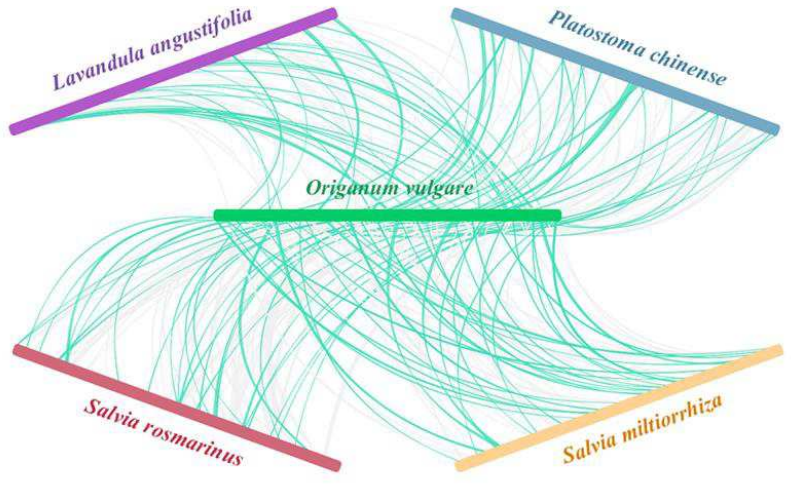
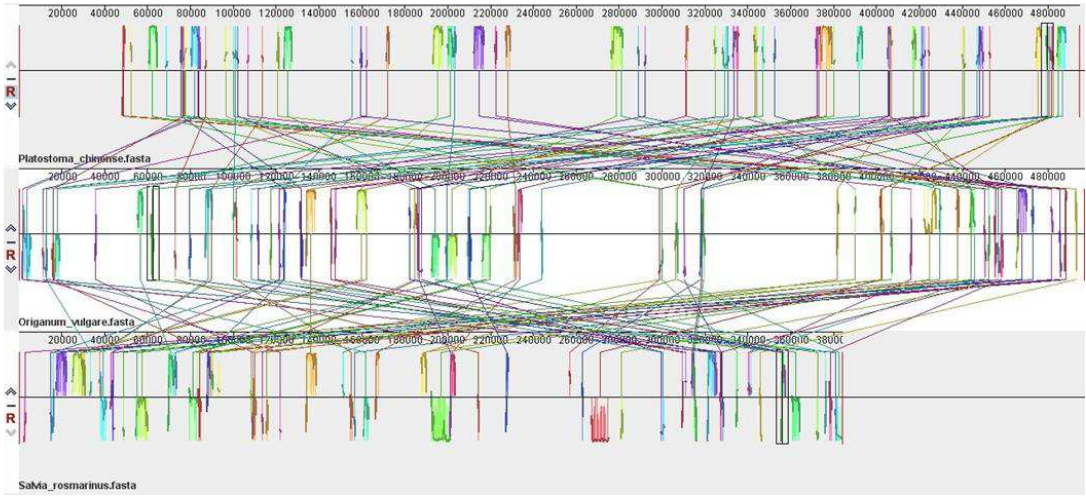


Figure 11 | Comparative Genomics Analysis of the mitogenome among *O.vulgare* and other relative species. (A) Grey areas indicate covariant blocks in the same direction, while green areas indicate covariant blocks where inversion occurs. (B) Mitogenome synteny between *P. chinense*, *O.vulgare* and *S. rosmarinus*.

A



B



978 **Table 1 | Encoding genes of *O. vulgare* mitogenome.**

Group of genes	Name of genes
Core genes	ATP synthase <i>atp1, atp4, atp6(x2), atp8,atp9</i>
	NADH dehydrogenase <i>nad1(x2), nad2, nad3, nad4, nad4L(x2), nad5, nad6, nad7(x2), nad9</i>
	Ubichinol cytochrome c reductase <i>cob(x2)</i>
	Cytochrome c biogenesis <i>ccmB, ccmC, ccmFC, ccmFN</i>
	Cytochrome c oxidase <i>cox1, cox2, cox3,</i>
	Maturases <i>matR</i>
Plastid-derived genes	Succinate dehydrogenase <i>sdh4</i>
	Large subunit of ribosome <i>rpl2, rpl5, rpl23</i>
	Small subunit of ribosome <i>rps4, rps12(x2), rps13</i>
rRNA genes	Ribosome RNA <i>rrn5(x2), rrn18, rrn26</i>
tRNA genes	Transfer RNA <i>trnA-UGC, trnfM-CAU, trnI-CAU(x3), trnL-CAA, trnL-UAG, trnM-CAU(x2), trnN-GUU, trnQ-UUG, trnS-UGA, trnW-CCA, trnY-GUA</i>

Note: Numbers in parentheses indicate gene copy counts.

981 **Table 2 | Primer sequences and PCR conditions for Verification of Homologous Recombination.**

Repeat	Primers	Product length (bp)	Sequence 5'-3'	PCR procedure	Refs
r227	F1	1337	GAACTCTTCCCGCTTCGACA	95°C 3 min; 94°C 15s, 53°C 30s, 72°C 1 min (35 cycles); 72°C 5 min	[17]
	R1		AAGATGTTTGGGCAGGGGAG		
	F2	657	TTCTGGTGTGGATGAGCGAC		
	R2		GATCGGGGTCATGCGTTTTG		
r287	F1	601	AGCTATCAAAGAGACGGGCG		
	R1		TAGAAGCCGGCGAAAGTGTT		
	F2	630	TTTATCTGCCCTAGCCCCCT		
	R2		GCTCTGCATCCCCTCAGTTT		
r652	F1	826	GAACTCTTCCCGCTTCGACA		
	R1		CTCCGATCCTCTCCGAGTCT		
	F2	657	TTCTGGTGTGGATGAGCGAC		
	R2		GATCGGGGTCATGCGTTTTG		
r510	F1	961	CCTACCCATGAAGCGCAGAT		
	R1		GACAGGGTGTGAATGCAGGA		
	F2	963	AAGAAACAGGCGTCCTCTCG		
	R2		TCTACGGTTCCATTGCGTGT		

P652	F1	883	GCTGCCTCCCTTTTCAATCC
	R1		CAGAGGTTGATGTCCCGGTG
	F2	1123	CGCCCATAGAGCTTAGCCAA
	R2		GGGCTAAAATCGACCCCTCT
P1553	F1	1957	ACTGGATTACCTGAAGGCGC
	R1		CCAATTAACCCTCCCCTCCG
	F2	1717	GGCTGGCTCCGTACTAATCC
	R2		CGTTCGATAGCTACCTGGGG
P144	F1	1145	CGGAGGGGCAAAGTCAATCT
	R1		CGACTTACCCGCCTTCTGTT
	F2	1857	GAGAGAGCGAGTGTCCGAAG
	R2		AGGAAGGAGGCACGCTACTA

Supplementary Files

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

- [FigureS1.pdf](#)
- [FigureS2.pdf](#)
- [FigureS3.pdf](#)
- [FigureS4.pdf](#)
- [FigureS5.pdf](#)
- [TableS1.docx](#)
- [TableS2.docx](#)
- [TableS3.docx](#)
- [TableS4.docx](#)
- [TableS5.docx](#)
- [TableS6.docx](#)
- [TableS7.docx](#)
- [TableS8.docx](#)
- [TableS9.docx](#)
- [TableS10.docx](#)
- [TableS11.docx](#)