

Chloroplast Genome Architecture of *Trichosanthes kirilowii* Maxim.: Structural Evolution, Repetitive Element Dynamics, and Comparative Phylogenetic Insights in Cucurbitaceae

Zhuanzhuan Jiang

zzjiang@whu.edu.cn

Anqing Normal University

Xinyi Ge

Anqing Normal University

Zehui Yang

Anqing Normal University

Ruyin Yuan

Anqing Normal University

Research Article

Keywords: *Trichosanthes kirilowii* Maxim, chloroplast genome, comparative genomics, phylogeny

Posted Date: June 3rd, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-6691477/v1>

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

[Read Full License](#)

Abstract

As a medicinally significant species within the Cucurbitaceae family, *Trichosanthes kirilowii* is currently lacking comprehensive chloroplast genomic resources that are essential for conducting evolutionary and phylogenetic studies. In this study, we present a characterization of its complete chloroplast genome, which consists of a circular DNA molecule measuring 156,924 bp with a quadripartite structure: LSC (85,994 bp), SSC (18,334 bp), and two IRs (52,558 bp each). This genome encompasses 132 functional genes and exhibits elevated GC content in the IRs (43.01%) compared to the LSC (34.89%) and SSC (31.3%). Analysis of repetitive sequences identified 198 long repeats as well as SSR motifs predominantly consisting of mononucleotide (A/T) variants. Codon usage displayed a strong preference for A/U-ending codons, revealing an extreme bias towards AGA and GAC. Comparative analysis across ten Cucurbitaceae species uncovered conserved IR/SC boundaries except in *Cucumis melo*—which exhibited divergent positioning of *rpl2*—and *Lagenaria siceraria*—which showed a complete loss of *ndhF*. Hypervariable regions such as *rps16* and *psbK-trnS* along with non-coding spacers like *trnQ-trnR* exhibited increased nucleotide diversity (Pi), contrasting sharply with the conservation observed in the IR regions. Evolutionary rate analysis indicated predominant purifying selection with Ka/Ks ratios ≤ 1 . Phylogenetic reconstruction successfully positioned *T. kirilowii* within a distinct evolutionary clade, thus providing valuable genomic foundations for molecular identification and biogeographic studies within the Cucurbitaceae family. This investigation establishes critical chloroplast genomic resources while also identifying mutation hotspots which can be explored further in future phylogenetic research endeavors.

Introduction

The chloroplast genome, an essential genetic resource in photosynthetic eukaryotes, serves as a cornerstone for understanding phylogenetic relationships and advancing plant systematics (Liu et al., 2024; Wang et al., 2024c). Its uniparental inheritance, conserved structure, and moderate evolutionary rate provide a solid molecular framework to address taxonomic uncertainties in morphologically complex plant groups (Xiao and Ge, 2022). Recent advancements in high-throughput sequencing have revolutionized comparative chloroplast genomics, allowing systematic exploration of structural variations, nucleotide diversity, and evolutionary constraints across plant lineages (Mahai et al., 2024; Yang et al., 2024). These developments are particularly impactful for elucidating intricate phylogenetic relationships within challenging taxonomic families that often blur traditional classification due to convergent morphological traits or incomplete lineage sorting (Ma et al., 2024; Sharko et al., 2024; Tamboli et al., 2025). In this context, chloroplast genomics has proven to be instrumental in elucidating evolutionary trajectories and refining taxonomic boundaries. The canonical quadripartite structure of chloroplast genomes—consisting of large and small single-copy regions (LSC/SSC) flanked by inverted repeats (IRs)—serves as a molecular framework for comparative analyses. While IR regions typically exhibit structural conservation due to mechanisms of copy correction, the LSC and SSC regions demonstrate accelerated rates of nucleotide substitution, making them particularly informative for phylogenetic reconstruction at lower taxonomic levels (Liu et al., 2024; Yang et al., 2023). Furthermore,

the identification of hypervariable loci, biases in codon usage, and architectures of repetitive sequences within chloroplast genomes has facilitated the development of novel molecular markers for species authentication and population genetics (Khan et al., 2021; Wen et al., 2021). Collectively, these genomic features provide critical insights into plant adaptation mechanisms, lineage-specific evolutionary pressures, and the molecular basis underlying photosynthetic efficiency—a trait that is paramount in both natural ecosystems and cultivated species (Ye et al., 2024).

The genus *Trichosanthes* (Cucurbitaceae), comprising approximately 100 species distributed across tropical and subtropical Asia, holds significant relevance in the current genomic era (Guo et al., 2020; Ma et al., 2022). Among these, *Trichosanthes kirilowii* Maxim., commonly referred to as Gualou, is distinguished as a species of exceptional pharmacological and economic importance (Liu et al., 2022; Song et al., 2023). For over two millennia, this perennial vine has been integral to traditional Chinese medicine (TCM), with its dried roots, fruits, and seeds extensively utilized for the treatment of inflammatory disorders, cardiovascular diseases, and respiratory conditions (Lai et al., 2024; Pabuprapap and Suksamrarn, 2021; Song et al., 2025). Modern phytochemical analyses have identified triterpenoid saponins, trichosanthin proteins, and flavonoid glycosides as key bioactive constituents responsible for its documented antitumor, antioxidant, and hypoglycemic properties (Chen et al., 2018a; Lu et al., 2022; Xu et al., 2012). Beyond therapeutic applications, the commercial value of *T. kirilowii* extends to its starch-rich tubers which serve as nutritional supplements; additionally, seed oils demonstrate potential for cosmetic and industrial uses (Wang et al., 2024a; Wang et al., 2024b). However, the sustainable utilization of *T. kirilowii* resources confronts critical challenges including taxonomic confusion with morphologically similar congeners as well as inadequate genomic tools for authenticating medicinal materials—issues that directly impede product quality control and intellectual property protection within herbal markets.

Despite these advancements, significant knowledge gaps remain concerning the chloroplast genomic landscape of *Trichosanthes* and its phylogenetic relatives. Most existing genomic resources within the Cucurbitaceae family predominantly focus on economically important taxa such as *Cucumis* (cucumbers) and *Citrullus* (watermelons) (Kousar and Park, 2023; Li et al., 2019; Rodríguez-Moreno et al., 2011). This has resulted in medicinally relevant genera like *Trichosanthes* being underrepresented in comparative genomic databases. Such disparities impede a comprehensive understanding of evolutionary patterns across the family and restrict the development of species-specific molecular tools for quality control purposes in herbal industries. Moreover, the evolutionary forces influencing codon usage preferences, repetitive sequence distributions, and selective constraints within Cucurbitaceae chloroplast genomes are still poorly characterized—gaps that limit our capacity to interpret functional correlations between genomic features and phenotypic traits (Ma et al., 2022).

This study establishes foundational knowledge by presenting the first comprehensive chloroplast genomic analysis of *T. kirilowii* cultivated variety "Wanlou 20", employing comparative genomic approaches to address longstanding taxonomic and evolutionary questions in Cucurbitaceae. Through meticulous examination of genomic architecture, molecular evolution patterns, and phylogenetic signals,

we aim to provide critical insights that bridge the gap between traditional plant systematics and modern genomic science, ultimately supporting the sustainable utilization of this pharmacologically vital species.

Materials and methods

Plant materials and growth condition

The mature seeds of *T. kirilowii* were sourced from the Provincial Key Laboratory of Biodiversity Study and Ecological Conservation in Southwest Anhui Province. Following a germination protocol, seeds underwent 24-hour aqueous imbibition at room temperature before being placed in a controlled-environment growth chamber (28°C, 60% relative humidity, darkness). Germinated seedlings with radical emergence were then transplanted into a nutrient-rich potting mix and grown in a climate-controlled greenhouse under standardized conditions: 16-hour photoperiod followed by an 8-hour dark cycle, with day temperatures regulated at 25°C and relative humidity maintained at 65–70%.

DNA extraction and sequencing

The chloroplast genomic DNA was isolated from 14-day-old juvenile foliage using an optimized CTAB-based extraction protocol (Murray and Thompson, 1980). Nucleic acid integrity was systematically validated through three complementary approaches: (1) electrophoretic separation on 1% agarose gel with ethidium bromide staining, (2) spectrophotometric quantification using a NanoPhotometer® N60 system (IMPLEN, USA) to assess A260/A280 purity ratios. To achieve comprehensive chloroplast genome assembly, we employed a dual-platform sequencing strategy that combined short-read and long-read technologies. Illumina sequencing was conducted on the HiSeq2500 platform (Illumina Inc., USA) with paired-end 150 bp reads generated from 300 bp insert-size libraries. Raw data underwent strict quality control using SOAPnuke (v2.1.4) (Chen et al., 2018b), which included removing reads with > 5% N bases, eliminating adapter contamination, and trimming low-quality bases (Q-score < 20). Long-read sequencing was performed on Oxford Nanopore Technology's PromethION platform, utilizing ligation sequencing kit according to manufacturer specifications for library preparation. Size-selected DNA fragments (≥ 10 kb) were loaded onto R9.4.1 Flow Cells (FLO-PRO002M) and sequenced for 48–72 hours in continuous mode. Real-time base calling and quality filtering were done using Guppy (v6.4.2) with the high-accuracy model (Wick et al., 2019).

Assembly and annotation of *T. kirilowii*

By employing a hybrid sequencing strategy that integrates the GetOrganelle bioinformatics pipeline with second-generation Illumina short reads and third-generation Oxford Nanopore long reads, accurate and contiguous assemblies of *T. kirilowii* have been achieved successfully. The chloroplast genome assembly began with GetOrganelle v1.7.7.0 (Jin et al., 2020) for extracting organelle-specific sequences from Illumina paired-end data. This produced preliminary unitig graphs that underwent quality control via

Bandage v0.8.1 (Wick et al., 2015), systematically identifying and excising mitochondrial and nuclear-derived unitigs based on topological features and read coverage patterns. Following this, corrected third-generation sequencing reads were aligned to the graph using minimap2 (Li, 2018), facilitating the extraction of chloroplast-derived long reads. The repeats present on the graph were resolved through alignment with these long reads, ultimately resulting in the construction of a circular DNA molecule. Finally, Geneious Prime (Kearse et al., 2012) was employed to verify assembly results by mapping Illumina reads onto the assembled chloroplast genome. The complete sequences of chloroplast genomes underwent annotation via CPGAVAS2 (Liu et al., 2023), followed by tRNAscan-SE (Chan et al., 2021) and HS-BLASTN (Chen et al., 2015) for tRNA and rRNA annotation respectively. Apollo v2.6.6 (Lewis et al., 2002) was then used to individually correct any annotation errors found in these chloroplast genomes. Illustrations of these genomes were generated using OGDRAW software (<https://chlorobox.mpimp.golm.mpg.de/OGDraw.html>).

Analysis of repetitive sequences in chloroplast genomes

The chloroplast genome's repetitive sequences (including forward, palindromic, complementary, and reverse repeats) were characterized using REPuter (<https://bibiserv.cebitec.uni-bielefeld.de/reputer/>) (Kurtz et al., 2001). The analysis parameters included: (1) Hamming distance of 3; (2) minimum repeat unit size of 12 bp; (3) maximum repeat length threshold of 50 bp. Simple sequence repeats (SSRs) in the organelle genome were identified using MISA software (version 2.1) (Beier et al., 2017), adhering to organelle genome analysis standards: mononucleotide repeats required a minimum of 10 iterations, dinucleotide motifs needed at least 6 units, while trinucleotide to hexanucleotide motifs had a threshold of 5 units. To ensure independent locus identification, an inter-SSR distance constraint of 100 base pairs was applied between adjacent microsatellite regions.

Analysis of codon usage bias

Codon usage analysis of chloroplast protein-coding genes was performed utilizing CodonW 1.44 software in conjunction with the EMBOSS Explorer platform (<http://bioinformatics.nl/emboss-explorer>). Relative Synonymous Codon Usage (RSCU) values were calculated to evaluate codon usage bias. RSCU is defined as the ratio of the observed frequency of a specific codon to its expected frequency under uniform usage within synonymous groups.

Comparative analysis of the chloroplast genome

To elucidate the evolutionary divergence patterns in genomic architecture between *T. kirilowii* and phylogenetically representative taxa within Cucurbitaceae, we systematically retrieved whole-genome sequencing data from nine additional cucurbitaceous species archived in the GenBank database (<https://www.ncbi.nlm.nih.gov/>). These species include *Citrullus amarus* (NC_035974.1), *Cucurbita pepo* (NC_038229.1), *Cucumis melo* (QR643675.1), *Lagenaria siceraria* (NC_036808.1), *Luffa aegyptiaca*

(QR327040.1), *Momordica cochinchinensis* (NC_065200.1), *Sicyos angulatus* (NC_062884.1), *Thladiantha cordifolia* (QR668859.1), *Gynostemma burmanicum* (NC_036141.1). Utilizing the IRscope (Amiryousefi et al., 2018), we accurately delineated the junctional boundaries of inverted repeat (IR) expansion within the chloroplast genome through comparative genomic visualization. Sequence polymorphism was characterized using sliding window analysis (window size: 600 bp; step size: 200 bp) in DnaSP6. This included calculating nucleotide diversity (π) indices for components of the chloroplast genome (LSC, SSC, and IR regions). Phylogenetic analysis of sequence divergence was conducted using DnaSP6 (version 61230) (Rozas et al., 2017) to assess evolutionary selection pressures by calculating nonsynonymous (K_a) and synonymous (K_s) substitution rates. K_a/K_s ratios were deemed biologically irrelevant when K_s values neared zero due to mathematical indeterminacy. Multispecies comparative analyses were visualized with heat maps generated in GraphPad Prism 8.0, followed by vector graphic optimization in Adobe Illustrator CC 2016 for publication-quality resolution.

Phylogenetic analysis

To elucidate the phylogenetic relationships of *T. kirilowii* within the order Cucurbitales, we conducted a comprehensive phylogenetic reconstruction through the comparative analysis of chloroplast genomic sequences from 40 representative species. The complete chloroplast genome sequences were subjected to extensive multiple sequence alignment using Mulan (<https://mulan.dcode.org/>). The resultant alignment file was subsequently utilized for phylogenetic tree construction in MEGA11 (Tamura et al., 2021), employing maximum likelihood methodology with 1000 bootstrap replicates to assess nodal support.

Result

Organization of Chloroplast Genome

The complete chloroplast genome of *T. kirilowii* measured 156,924 base pairs (bp) (Fig. 1). This chloroplast genome exhibited a typical tetrad structure and comprised a large single copy (LSC) region of 85,994 bp, a small single copy (SSC) region of 18,334 bp, and two inverted repeat regions (IRa and IRb), each measuring 52,558 bp. Notably, the genome showcased a high guanine-cytosine (GC) content of 37.1%, with specific GC contents of 34.89% for the LSC region, 31.3% for the SSC region, and an elevated value of 43.01% for the IR regions. The genome contained a total of 132 genes, which included 88 protein-coding genes, 36 transfer RNA genes, and eight ribosomal RNA genes (Table 1, Table 2).

Table 1
Chloroplast genome features of *T. kirilowii*

<i>T. kirilowii</i>	Length (bp)		GC content (%)	Gene number	
	genome	156,924		total genes	132
	LSC	85,994		protein coding genes	88
	SSC	18,334		rRNAs	8
	IR	52,558		tRNAs	36

Table 2
Genes within the chloroplast genome of *T. kirilowii*

Category for genes	Group of genes	Name of genes
Transcription and translation	Transfer RNAs	<i>trnA</i> -UGC(×2), <i>trnC</i> -GCA, <i>trnD</i> -GUC, <i>trnE</i> -UUC, <i>trnF</i> -GAA, <i>trnG</i> -UCC, <i>trnG</i> -GCC, <i>trnH</i> -GUG, <i>trnI</i> -CAU(×2), <i>trnI</i> -GAU(×2), <i>trnK</i> -UUU, <i>trnL</i> -CAA(×2), <i>trnL</i> -UAG, <i>trnL</i> -UAA, <i>trnM</i> -CAU, <i>trnN</i> -GUU(×2), <i>trnP</i> -UGG, <i>trnQ</i> -UUG, <i>trnR</i> -UCU, <i>trnR</i> -ACG(×2), <i>trnS</i> -GCU(×2), <i>trnS</i> -UGA, <i>trnT</i> -GGU, <i>trnT</i> -UGU, <i>trnV</i> -GAC(×2), <i>trnV</i> -UAC, <i>trnW</i> -CCA, <i>trnY</i> -GUA
	RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> , <i>rpoC2</i>
	Ribosomal proteins (SSU)	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> (×2), <i>rps8</i> , <i>rps11</i> , <i>rps12</i> (×3), <i>rps14</i> , <i>rps15</i> , <i>rps16</i> , <i>rps18</i> , <i>rps19</i>
	Ribosomal proteins (LSU)	<i>rpI2</i> (×2), <i>rpI14</i> , <i>rpI16</i> , <i>rpI20</i> , <i>rpI22</i> , <i>rpI23</i> (×2), <i>rpI32</i> , <i>rpI33</i> , <i>rpI36</i>
	Ribosomal RNAs	<i>rrn4.5S</i> (×2), <i>rrn5S</i> (×2), <i>rrn16S</i> (×2), <i>rrn23S</i> (×2)
Photosynthesis	NADH dehydrogenase	<i>ndhA</i> , <i>ndhB</i> (×2), <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
	Photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaJ</i> , <i>psaI</i>
	Photosystem II	<i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i>
	Cytochrome b/f complex	<i>petA</i> , <i>petB</i> , <i>petD</i> , <i>petG</i> , <i>petL</i> , <i>petN</i>
	Rubisco	<i>rbcL</i>
	ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> , <i>atpH</i> , <i>atpI</i>
Other genes	Maturase	<i>matK</i>
	Protease	<i>clpP</i>
	Envelope membrane protein	<i>cemA</i>
	Subunit of Acetyl-CoA carboxylase	<i>accD</i>
	C-type cytochrome synthesis gene	<i>ccsA</i>
Unknown function	Hypothetical chloroplast	<i>ycf1</i> (×2), <i>ycf2</i> (×2), <i>ycf3</i> , <i>ycf4</i> , <i>ycf15</i> (×2)

Repetitive Sequence and Codon Usage

The chloroplast genome of *T. kirilowii* was found to contain 198 long repetitive elements, which were systematically classified into four distinct categories based on their orientation patterns. Specifically, these included forward repeats (67), palindromic repeats (67), reverse repeats (44), and complementary repeats (20), with forward-type repeats constituting the predominant category. The analysis of long repeat elements revealed a predominance of 18-nucleotide repeat sequences (29), encompassing all four fundamental types of repeats: forward, palindromic, reverse, and complementary. This was followed by 23-nucleotide repeat sequences (26), which similarly exhibited comprehensive representation across the four canonical categories of repeats (Fig. 2A).

Through comprehensive analysis of the *T. kirilowii* chloroplast genome, four distinct classes of simple sequence repeats (SSRs) were characterized based on their repeat unit lengths: mononucleotide (A/T), dinucleotide (AT/TA), tetranucleotide (AAAT/TAAA), and pentanucleotide (TAAAC) motifs. Notably, mononucleotide repeats exhibited the highest prevalence among the identified SSRs, demonstrating significantly greater abundance compared to the less frequent dinucleotide, tetranucleotide, and pentanucleotide variants (Fig. 2B).

The chloroplast genome of *T. kirilowii* comprises 88 protein-coding genes. Codon usage bias analysis revealed a pronounced preference for 3'-terminal A/U-ending codons, with the majority demonstrating relative synonymous codon usage (RSCU) values ≥ 1 . Conversely, codons terminating in C/G at the third position exhibited reduced RSCU values, predominantly < 1 . Notably, AGA (Arginine) emerged as the most prevalent codon (RSCU = 1.88), whereas GAC (Aspartate) displayed the lowest utilization frequency (RSCU = 0.54). Intriguingly, two amino acids - Tryptophan (UGG) and Methionine (AUG) - exhibited no codon usage bias, both maintaining RSCU values of precisely 1.00 (Fig. 3).

Expansion and Contraction of Inverted Repeats in Chloroplast Genomes

Boundary expansion analysis across ten Cucurbitaceae species demonstrated conserved IR/SC junction patterns with notable exceptions. The LSC/IRb junction was localized between *rps19* and *rpl2* in nine species, while exhibiting a distinct displacement to the *rpl2* locus in *C. melo*. Comparative analysis revealed that the IRb/SSC boundary was embedded within the *ndhF* coding region in nine taxa, with the exception of *L. siceraria* which showed complete loss of this gene. Regarding the SSC/IRa junction, eight species displayed characteristic positioning within *ycf1*, contrasting with *M. cochinchinensis* and *C. pepo* that exhibited divergent boundary configurations. Furthermore, phylogenetic conservation was observed at the IRa/LSC interface where eight species shared the canonical *rpl2-tRNA-His* boundary organization, whereas *C. melo* and *C. pepo* demonstrated evolutionary divergence through alternative junction placements (Fig. 4).

Comparative Analysis of the Chloroplast Genome Sequences in Cucurbitaceae

Comparative analysis of chloroplast genome nucleotide diversity (π) revealed distinct patterns of sequence divergence across genomic compartments (Fig. 5). The LSC and SSC regions exhibited elevated nucleotide diversity compared to inverted repeat regions, with particularly pronounced variation observed in eight hypervariable loci: *rps16*, *atpE-atpB*, *trnA-UCC*, *rrn16s*, *psbK-trnS-GCU*, *rps14-psaB*, *ndhI-ndhA*, and *trnI-GCU*. These molecular hotspots demonstrate accelerated evolutionary rates within Cucurbitaceae chloroplast genomes, suggesting they may represent critical targets for phylogenetic reconstruction and molecular marker development.

A comprehensive comparative analysis of chloroplast genome architecture across ten phylogenetically diverse Cucurbitaceae taxa was performed using mVISTA visualization, employing the fully annotated *T. kirilowii* chloroplast genome as the reference framework (Fig. 6). Our whole-genome alignment revealed distinct patterns of molecular evolution, with both LSC and SSC regions demonstrating significantly elevated sequence divergence relative to the conserved IR regions. Critical examination identified pronounced sequence polymorphisms in functionally important loci, including ribosomal protein-coding genes (*rps16*, *rps32*), NADH dehydrogenase subunits (*ndhA*), and the *ycf1* gene - a recognized hotspot for chloroplast genome evolution. Furthermore, non-coding regions exhibited substantial nucleotide diversity, particularly in key regulatory spacers: *trnQ-trnR*, *trnC-psbM*, *trnL-trnF*, and *rpl32-trnL*, suggesting potential evolutionary constraints on transcriptional regulation mechanisms. Comparative evolutionary analysis of chloroplast protein-coding loci across a phylogenetically diverse panel of 10 Cucurbitaceae species revealed that the majority of K_a/K_s ratios consistently exhibited values ≤ 1 . These molecular evolutionary patterns indicate that the plastid genomes have been predominantly governed by purifying selection with sporadic neutral evolutionary processes, suggesting strong functional constraints on photosynthetic apparatus genes and housekeeping genetic elements in chloroplast genomes.

Phylogenetic Relationship Reconstruction

To clarify the precise phylogenetic placement of *T. kirilowii* within the Cucurbitaceae clade, we conducted a maximum-likelihood phylogenetic reconstruction utilizing complete chloroplast genome sequences from several taxonomically representative species. In this analysis, *Arabidopsis thaliana* was designated as the outgroup taxon for establishing phylogenetic rooting. The 36 species were classified into two distinct branches: a larger branch and a smaller branch. The larger branch encompassed 29 species, while the smaller branch comprised 7 species. *T. kirilowii* was situated within the larger branch, positioned between loofah and *Sicyos angulatus*, thereby indicating their closer phylogenetic relationship (Fig. 8).

Discussion

Structural Conservation and Divergence in Chloroplast Genome Organization

The quadripartite structure of the *T. kirilowii* chloroplast genome (156,924 bp) aligns with the canonical organization observed in most angiosperms, comprising LSC, SSC, and IR regions (Fig. 1) (Tamboli et al., 2025). The elevated GC content in IR regions (43.01%) compared to single-copy regions (LSC: 34.89%; SSC: 31.3%) is consistent with patterns reported across plant lineages (Table 1), likely reflecting selection for stable secondary structures in rRNA genes and replication origins within IRs. Notably, the GC content gradient across genomic compartments mirrors findings in *Cucumis* species, suggesting a conserved mechanism for maintaining transcriptional efficiency in IR-located ribosomal operons (Yu et al., 2021). The gene repertoire of 132 genes, including 88 protein-coding genes (Table 2), aligns with the reduced plastome complexity typical of autotrophic Cucurbitaceae, contrasting with the gene loss patterns observed in parasitic plants. The retention of all *ndh* genes in *T. kirilowii* contrasts with their frequent pseudogenization in shade-adapted species, supporting its evolutionary trajectory as a sun-adapted lineage (Amar, 2020).

Repetitive Elements and Genome Plasticity

The identification of 198 long repetitive sequences, predominantly forward and palindromic repeats, underscores the role of repeat-mediated recombination in plastome evolution (Fig. 2). The abundance of 18- and 23-nt repeats suggests these elements may serve as recombination hotspots, potentially driving the structural variations observed in IR boundary shifts among Cucurbitaceae species. The prevalence of mononucleotide SSRs (A/T) over higher-order repeats aligns with replication slippage mechanisms dominating plastome microsatellite generation. These SSRs, particularly in non-coding regions, represent valuable molecular markers for population-level studies, as demonstrated in *Citrullus* species (Ebert and Peakall, 2009). The codon usage bias toward A/U-ending codons ($\text{RSCU} \geq 1$) likely reflects co-evolution with the plastid translational machinery, optimizing translation efficiency through tRNA abundance matching—a phenomenon well-documented in *Arabidopsis* and *Oryza* (Kawabe and Miyashita, 2003; Yang and Nielsen, 2000). The exceptional neutrality of tryptophan (UGG) and methionine (AUG) codons (Fig. 3), constrained by their single-codon assignments, highlights the interplay between genetic code degeneracy and selective pressures (Morton, 1998).

Dynamic Evolution of Inverted Repeat Boundaries

The IR boundary analysis across ten Cucurbitaceae species revealed both conserved features and lineage-specific innovations. The displacement of the LSC/IRb junction to *rpl2* in *Cucumis melo* and the complete loss of *ndhF* in *Lagenaria siceraria* exemplify the evolutionary plasticity of IR/SC junctions (Fig. 4) (Zhang et al., 2018). These boundary shifts likely result from homologous recombination between small repeats near IR ends, as proposed in *Nicotiana* (Wang et al., 2022). The retention of *ycf1* spanning the SSC/IRa boundary in most species supports its hypothesized role in IR stabilization, while its absence in *Momordica cochinchinensis* and *Cucurbita pepo* suggests alternative stabilization mechanisms. The phylogenetic conservation of the *rpl2*-tRNA-His boundary at IRa/LSC in eight species contrasts with divergent configurations in *C. melo* and *C. pepo*, potentially reflecting differential selection on DNA repair pathways among lineages (Sun et al., 2018).

Hypervariable Regions and Molecular Evolution

The elevated nucleotide diversity (π) in LSC/SSC regions versus IRs corroborates the paradigm of accelerated evolution in single-copy regions due to reduced selective constraints on non-coding spacers (Fig. 5) (Wicke et al., 2011). The identification of eight hypervariable loci (*rps16*, *atpE-atpB*, *trnA-UCC*, etc.) (Fig. 6) expanded the toolkit for Cucurbitaceae phylogenetics, with *trnL* intron and *rpl20-rps12* spacer previously validated as markers in *Cucumis* (Sebastian et al., 2010). The high divergence in *ycf1* and *ndhA* aligns with their roles as evolutionary hotspots in angiosperm plastomes, potentially linked to their large size and presence in recombination-prone regions (Jansen et al., 2007). The conserved IR regions, housing ribosomal RNA operons, exhibit strong purifying selection, consistent with functional constraints on photosynthetic and translational apparatuses. However, the elevated diversity in intergenic spacers (*trnQ-trnR*, *rpl32-trnL*) suggests relaxed selection in regulatory regions, possibly permitting cis-regulatory evolution without compromising protein function (Daniell et al., 2016).

Phylogenetic Implications

The observed prevalence of K_a/K_s ratios ≤ 1 across chloroplast protein-coding loci in Cucurbitaceae species highlighted the significant role of purifying selection in plastid genome evolution (Fig. 7). This aligns with evolutionary theory, as purifying selection preserves essential functional elements, especially in photosynthesis-related genes and housekeeping machinery (Gandullo et al., 2019). Such constraints reflect the critical function of chloroplast-encoded proteins in maintaining photosynthetic efficiency and cellular homeostasis, allowing little room for nonsynonymous substitutions that could impair protein function. These findings are consistent with patterns seen in other angiosperm lineages like Malvaceae, where chloroplast genomes also show strong purifying selection alongside localized neutral drift (Ruang-areerate et al., 2025).

Phylogenetic reconstruction clearly placed *T. kirilowii* within the Cucurbitaceae clade, situated between loofah and *Sicyos angulatus*. This placement supported recent molecular systematics that redefine Cucurbitaceae taxonomy by grouping Trichosanthes with core clade members sharing floral synapomorphies and fruit developmental pathways (Calvo et al., 2013). However, its closer relationship to loofah than to Momordica or Citrullus contrasts with earlier morphology-based classifications, suggesting convergent evolution of fruit traits within Cucurbitaceae (Guo et al., 2020).

Conclusion

This study deciphers the architectural blueprint and evolutionary forces shaping the *T. kirilowii* chloroplast genome, establishing a foundation for comparative genomics in Cucurbitaceae. The identified hypervariable regions and SSRs provide practical tools for phylogenetics and population genetics, while the IR boundary analysis advances our understanding of plastome stability mechanisms. By bridging structural genomics and molecular evolution, this work exemplifies how chloroplast genome sequencing can address taxonomic challenges and illuminate evolutionary processes in plant families with ecological and economic significance.

Declarations

Author Contributions

The contributions of the authors are organized as follows: Zhuanzhuan Jiang conceived the study, developed software, validated methodologies, conducted investigations, acquired resources, curated data, drafted the initial manuscript, and contributed to both visualization and funding acquisition. Xinyi Ge participated in conceptual design, performed formal analyses, engaged in manuscript review and editing, visualization efforts, and provided funding support. Zehui Yang contributed to conceptualization processes, formal analytical activities, critical revisions of the manuscript content, visual presentations, and procurement of financial resources. Ruyin Yuan assisted in the development of the conceptual framework as well as data curation and funding acquisition. All coauthors have rigorously evaluated the content of the manuscript and granted formal endorsement for its publication.

Data Availability

The data presented in this study are openly accessible in the NCBI GenBank (<https://www.ncbi.nlm.nih.gov/>) under accession number PV426762.

Funding

This work was funded by the State Key Laboratory of Hybrid Rice Program at Wuhan University (KF202204), the Scientific Research Foundation for Higher Education Institutions in Anhui Province (2022AH051034, 2022AH051048), the University Excellent Talent Training Project for 2023 (JWFX2023027), and the Key Laboratory of Biodiversity Conservation and Characteristic Resource Utilization in Southwest Anhui (Wxn202411).

Conflict of interest

The authors declare no conflict of interest.

Ethical approval

Not Applicable.

Consent to participate

Not Applicable.

Consent for publication

Not Applicable.

References

1. Amar, M.H. (2020). *ycf1-ndhF* genes, the most promising plastid genomic barcode, sheds light on phylogeny at low taxonomic levels in *Prunus persica*. *Journal, genetic engineering & biotechnology* 18:42.
2. Amiryousefi, A., Hyvonen, J., and Poczai, P. (2018). IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* 34:3030-3031.
3. Beier, S., Thiel, T., Munch, T., Scholz, U., and Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33:2583-2585.
4. Calvo, J., Álvarez, I., Aedo, C., and Pelter, P.B. (2013). A phylogenetic analysis and new delimitation of *Senecio* sect. *Crociseris* (Compositae: Senecioneae), with evidence of intergeneric hybridization. *Taxon* 62:127-140.
5. Chan, P.P., Lin, B.Y., Mak, A.J., and Lowe, T.M. (2021). tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic acids research* 49:9077-9096.
6. Chen, W., Zhang, S., You, J., Chen, Z., and Yue, M. (2018a). Analysis of Fatty Acids in *Trichosanthes kirilowii* Maxim by MSPD and Fluorescence Derivatization. *Journal of chromatographic science* 56:941-947.
7. Chen, Y., Chen, Y., Shi, C., Huang, Z., Zhang, Y., Li, S., Li, Y., Ye, J., Yu, C., Li, Z., et al. (2018b). SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience* 7:1-6.
8. Chen, Y., Ye, W., Zhang, Y., and Xu, Y. (2015). High speed BLASTN: an accelerated MegaBLAST search tool. *Nucleic acids research* 43:7762-7768.
9. Daniell, H., Lin, C.S., Yu, M., and Chang, W.J. (2016). Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome biology* 17:134-163.
10. Ebert, D., and Peakall, R. (2009). Chloroplast simple sequence repeats (cpSSRs): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. *Mol Ecol Resour* 9:673-690.
11. Gandullo, J., Monreal, J.A., Alvarez, R., Diaz, I., Garcia-Maurino, S., and Echevarria, C. (2019). Anionic Phospholipids Induce Conformational Changes in Phosphoenolpyruvate Carboxylase to Increase Sensitivity to Cathepsin Proteases. *Frontiers in plant science* 10:582.
12. Guo, J., Xu, W., Hu, Y., Huang, J., Zhao, Y., Zhang, L., Huang, C.-H., and Ma, H. (2020). Phylotranscriptomics in Cucurbitaceae Reveal Multiple Whole-Genome Duplications and Key Morphological and Molecular Innovations. *Molecular plant* 13:1117-1133.
13. Jansen, R.K., Cai, Z., Raubeson, L.A., Daniell, H., Depamphilis, C.W., Leebens-Mack, J., Muller, K.F., Guisinger-Bellian, M., Haberle, R.C., Hansen, A.K., et al. (2007). Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proceedings of the National Academy of Sciences of the United States of America* 104:19369-19374.
14. Jin, J.J., Yu, W.B., Yang, J.B., Song, Y., dePamphilis, C.W., Yi, T.S., and Li, D.Z. (2020). GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome biology*

15. Kawabe, A., and Miyashita, N.T. (2003). Patterns of codon usage bias in three dicot and four monocot plant species. . *Genes Genet Syst* 78:343-352.
16. Kears, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647-1649.
17. Khan, A.L., Asaf, S., Lubna, Al-Rawahi, A., and Al-Harrasi, A. (2021). Decoding first complete chloroplast genome of toothbrush tree (*Salvadora persica* L.): insight into genome evolution, sequence divergence and phylogenetic relationship within Brassicales. *BMC genomics* 22:312-328.
18. Kousar, M., and Park, J. (2023). Comparative Analysis of the Chloroplast Genome of *Sicyos angulatus* with Other Seven Species of Cucurbitaceae Family. *Genes* 14:1-12.
19. Kurtz, S., Choudhuri, J.V., Ohlebusch, E., Schleiermacher, C., J, S., and R, G. (2001). REPuter—the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research*, 29:4633-4642.
20. Lai, H.-C., Cheng, J.-C., Yip, H.-T., Jeng, L.-B., and Huang, S.-T. (2024). Chinese herbal medicine decreases incidence of hepatocellular carcinoma in diabetes mellitus patients with regular insulin management. *World Journal of Gastrointestinal Oncology* 16:716-731.
21. Lewis, S., Searle, S., Harris, N., Gibson, M., Iyer, V., Richter, J., Wiel, C., Bayraktaroglu, L., Birney, E., Crosby, M., et al. (2002). Apollo—a sequence annotation editor. *Genome biology* 3:1-14.
22. Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094-3100.
23. Li, Q., Li, H., Huang, W., Xu, Y., Zhou, Q., Wang, S., Ruan, J., Huang, S., and Zhang, Z. (2019). A chromosome-scale genome assembly of cucumber (*Cucumis sativus* L.). *GigaScience* 8:1-10.
24. Liu, Q., Dai, J., Chen, J., Liu, Z., Lin, Y., Qiu, G., Gao, X., Zhang, R., and Zhu, S. (2024). Comparative analysis the chloroplast genomes of *Celastrus* (Celastraceae) species: Provide insights into molecular evolution, species identification and phylogenetic relationships. *Phytomedicine : international journal of phytotherapy and phytopharmacology* 131:155770.
25. Liu, S., Ni, Y., Li, J., Zhang, X., Yang, H., Chen, H., and Liu, C. (2023). CPGView: A package for visualizing detailed chloroplast genome structures. *Molecular Ecology Resources* 23:694-704.
26. Liu, Y., Zhong, H., Xu, P., Zhou, A., Ding, L., Qiu, J., Wu, H., and Dai, M. (2022). Deciphering the combination mechanisms of Gualou–Xiebai herb pair against atherosclerosis by network pharmacology and HPLC-Q-TOF-MS technology. *Frontiers in Pharmacology* 13.
27. Lu, L., Peng, J., Wan, P., Peng, H., Lu, J., and Xiong, G. (2022). Mechanism of *Tripterygium wilfordii* Hook.F.- *Trichosanthes kirilowii* Maxim decoction in treatment of diabetic kidney disease based on network pharmacology and molecular docking. *Frontiers in Pharmacology* 13.
28. Ma, L., Wang, Q., Zheng, Y., Guo, J., Yuan, S., Fu, A., Bai, C., Zhao, X., Zheng, S., Wen, C., et al. (2022). Cucurbitaceae genome evolution, gene function, and molecular breeding. *Horticulture research* 9:1-24.

29. Ma, W.B., Ou, Y., Dayananda, B., Ji, H.J., and Yu, T. (2024). The complete chloroplast genome of *Rhododendron ambiguum* and comparative genomics of related species. *Comparative Cytogenetics* 18:143-159.
30. Mahai, R., Sheng, S., Wang, X., Yuan, J., and Mu, Z. (2024). Comparative analysis of complete chloroplast genomes of 14 Asteraceae species. *Molecular biology reports* 51:1094.
31. Morton, B.R. (1998). Selection on the Codon Bias of Chloroplast and Cyanelle Genes in Different Plant and Algal Lineages. *Journal of molecular evolution* 46:449-459.
32. Murray, M.G., and Thompson, W.F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic acids research* 8:4321-4325.
33. Pabuprapap, W., and Suksamrarn, A. (2021). Chemical constituents of the genus *Trichosanthes* (Cucurbitaceae) and their biological activities: A review. *ScienceAsia* 47S:1.
34. Rodríguez-Moreno, L., González, V.M., Benjak, A., Martí, M.C., Puigdomènech, P., Aranda, M.A., and Garcia-Mas, J. (2011). Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin. *BMC genomics* 12:1-14.
35. Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., and Sanchez-Gracia, A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Molecular biology and evolution* 34:3299-3302.
36. Ruang-areerate, P., Sangsrakru, D., Yoocha, T., Kongkachana, W., U-Thoomporn, S., Prathip Na Thalang, O., Chumriang, P., Wanthongchai, P., Tangphatsornruang, S., and Pootakham, W. (2025). Comparative Chloroplast Genomes and Phylogenetic Relationships of True Mangrove Species *Brownlowia tersa* and *Brownlowia argentata* (Malvaceae). *Current Issues in Molecular Biology* 47:74.
37. Sebastian, P., Schaefer, H., Telford, I.R., and Renner, S.S. (2010). Cucumber (*Cucumis sativus*) and melon (*C. melo*) have numerous wild relatives in Asia and Australia, and the sister species of melon is from Australia. *Proceedings of the National Academy of Sciences of the United States of America* 107:14269-14273.
38. Sharko, F.S., Petrova, K.O., Patrushev, M.V., Fedosov, D.Y., and Toshchakov, S.V. (2024). Chloroplast Genome Variation and Phylogenetic Relationships of Autochthonous Varieties of *Vitis vinifera* from the Don Valley. *International journal of molecular sciences* 25:9928.
39. Song, Q., Zhang, K., Gu, L., Zhang, C., and Bhuyar, P. (2023). Isolation, Purification, Structural Characterization, and Hypoglycemic Activity of Polysaccharide from *Trichosanthes kirilowii* Maxim. Seed Shell. *Journal of Food Processing and Preservation* 2023:1-13.
40. Song, Q., Zhang, K., Li, S., and Weng, S. (2025). *Trichosanthes kirilowii* Maxim. Polysaccharide attenuates diabetes through the synergistic impact of lipid metabolism and modulating gut microbiota. *Current Research in Food Science* 10:100977.
41. Sun, G., Xu, Y., Liu, H., Sun, T., Zhang, J., Hettenhausen, C., Shen, G., Qi, J., Qin, Y., Li, J., et al. (2018). Large-scale gene losses underlie the genome evolution of parasitic plant *Cuscuta australis*. *Nature*

communications 9:2683.

42. Tamboli, A.S., Youn, J.-S., Kadam, S.K., Pak, J.H., and Choo, Y.-S. (2025). Chloroplast Genome of *Arisaema takesimense*: Comparative Genomics and Phylogenetic Insights into the *Arisaema*. *Biochemical Genetics*:1-22.
43. Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular biology and evolution* 38:3022-3027.
44. Wang, S., Gao, J., Chao, H., Li, Z., Pu, W., Wang, Y., and Chen, M. (2022). Comparative Chloroplast Genomes of *Nicotiana* Species (Solanaceae): Insights Into the Genetic Variation, Phylogenetic Relationship, and Polyploid Speciation. *Frontiers in plant science* 13:899252.
45. Wang, W., Li, S., Zhu, Y., Cui, X., Sheng, Z., Wang, H., and Cheng, Z. (2024a). Antioxidant and Neuroprotective Effects of Seed Oils from *Trichosanthes kirilowii* and *T. laceribractea* in *Caenorhabditis elegans*: A Comparative Analysis and Mechanism Study. *Antioxidants* 13:861.
46. Wang, W., Li, S., Zhu, Y., Zhu, R., Du, X., Cui, X., Wang, H., and Cheng, Z. (2024b). Effect of Different Edible *Trichosanthes* Germplasm on Its Seed Oil to Enhance Antioxidant and Anti-Aging Activity in *Caenorhabditis elegans*. *Foods* 13:503.
47. Wang, X., Zhang, Z., Shi, Y., Man, J., Huang, Y., Zhang, X., Liu, S., He, G., An, K., Amu, L., et al. (2024c). Population identification and genetic diversity analysis of *Fritillaria ussuriensis* (*Fritillaria*) based on chloroplast genes *atpF* and *petB*. *Journal of applied genetics* 65:453-462.
48. Wen, F., Wu, X., Li, T., Jia, M., Liu, X., and Liao, L. (2021). The complete chloroplast genome of *Stauntonia chinensis* and compared analysis revealed adaptive evolution of subfamily *Lardizabaloideae* species in China. *BMC genomics* 22:1-18.
49. Wick, R.R., Judd, L.M., and Holt, K.E. (2019). Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome biology* 20:129.
50. Wick, R.R., Schultz, M.B., Zobel, J., and Holt, K.E. (2015). Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* 31:3350-3352.
51. Wicke, S., Schneeweiss, G.M., dePamphilis, C.W., Müller, K.F., and Quandt, D. (2011). The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant molecular biology* 76:273-297.
52. Xiao, T.W., and Ge, X.J. (2022). Plastome structure, phylogenomics, and divergence times of tribe *Cinnamomeae* (Lauraceae). *BMC genomics* 23:642.
53. Xu, Y., Chen, G., Lu, X., Li, Z.-Q., Su, S.-S., Zhou, C., and Pei, Y.-H. (2012). Chemical constituents from *Trichosanthes kirilowii* Maxim. *Biochemical Systematics and Ecology* 43:114-116.
54. Yang, L., Zhu, Y., and Hua, Q. (2024). Comparative chloroplast genomics of *Caryophyllaceae* species: insights into sequence variations and phylogenetic evolution. *BMC plant biology* 24:1259.
55. Yang, Q., Xin, C., Xiao, Q.S., Lin, Y.T., Li, L., and Zhao, J.L. (2023). Codon usage bias in chloroplast genes implicate adaptive evolution of four ginger species. *Frontiers in plant science* 14:1304264.

56. Yang, Z., and Nielsen, R. (2000). Estimating Synonymous and Nonsynonymous Substitution Rates Under Realistic Evolutionary Models. *Mol. Biol. Evol* 17:32-43.
57. Ye, J.-J., Lin, X.-Y., Yang, Z.-X., Wang, Y.-Q., Liang, Y.-R., Wang, K.-R., Lu, J.-L., Lu, P., and Zheng, X.-Q. (2024). The light-harvesting chlorophyll a/b-binding proteins of photosystem II family members are responsible for temperature sensitivity and leaf color phenotype in albino tea plant. *Journal of Advanced Research* 66:87-104.
58. Yu, Y., Li, H.T., Wu, Y.H., and Li, D.Z. (2021). Correlation Analysis Reveals an Important Role of GC Content in Accumulation of Deletion Mutations in the Coding Region of Angiosperm Plastomes. *Journal of molecular evolution* 89:73-80.
59. Zhang, X., Zhou, T., Yang, J., Sun, J., Ju, M., Zhao, Y., and Zhao, G. (2018). Comparative Analyses of Chloroplast Genomes of Cucurbitaceae Species: Lights into Selective Pressures and Phylogenetic Relationships. *Molecules* 23.

Figures

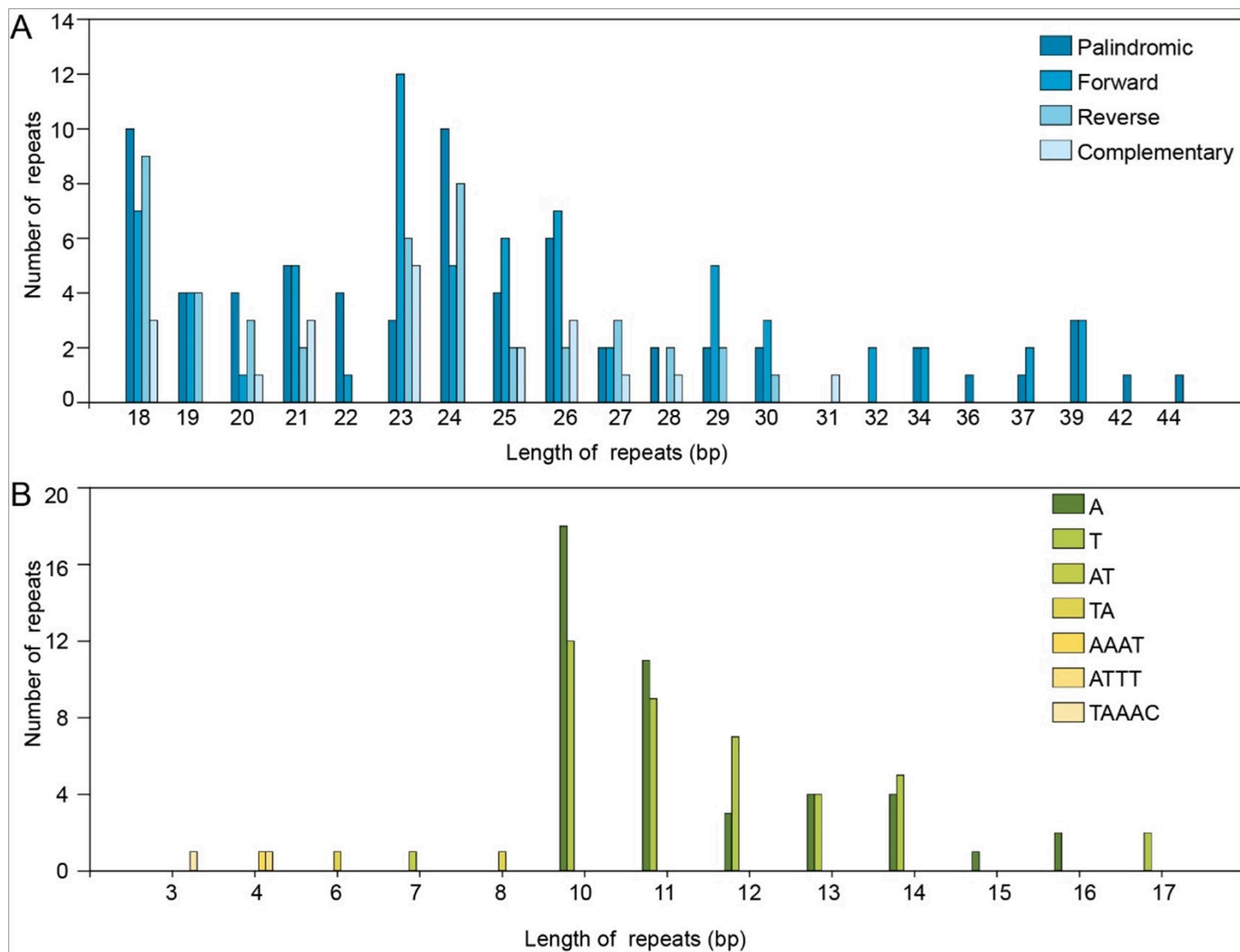


Figure 2

Repetitive sequences and SSR counts in the complete chloroplast genome of *T. kirilowii*

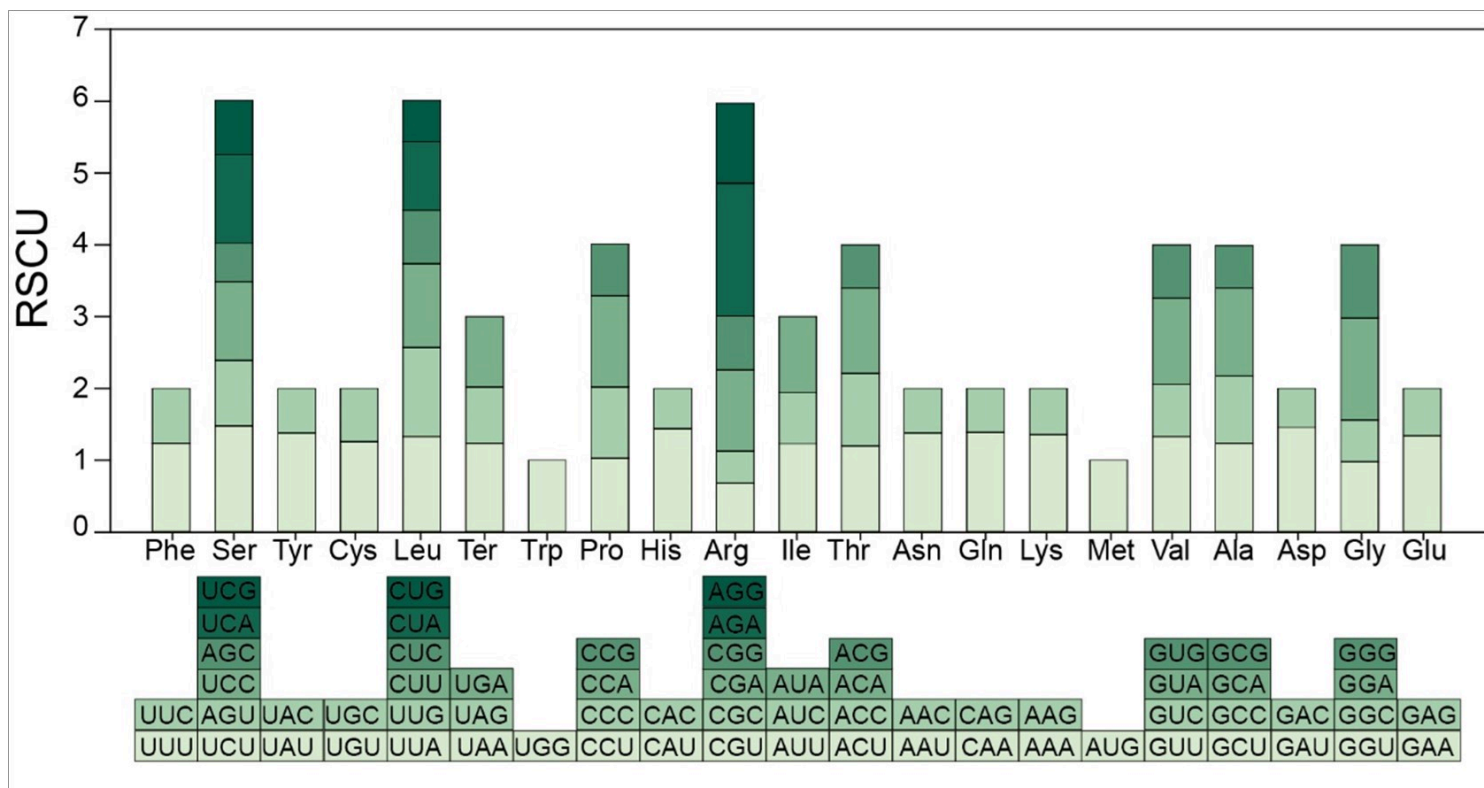


Figure 3

RSCU values for 21 amino acids in the *T. kirilowii* chloroplast genome CDS. The boxes below the graphs show all codons for each amino acid, with histogram colors matching those of the codons.

Inverted Repeats

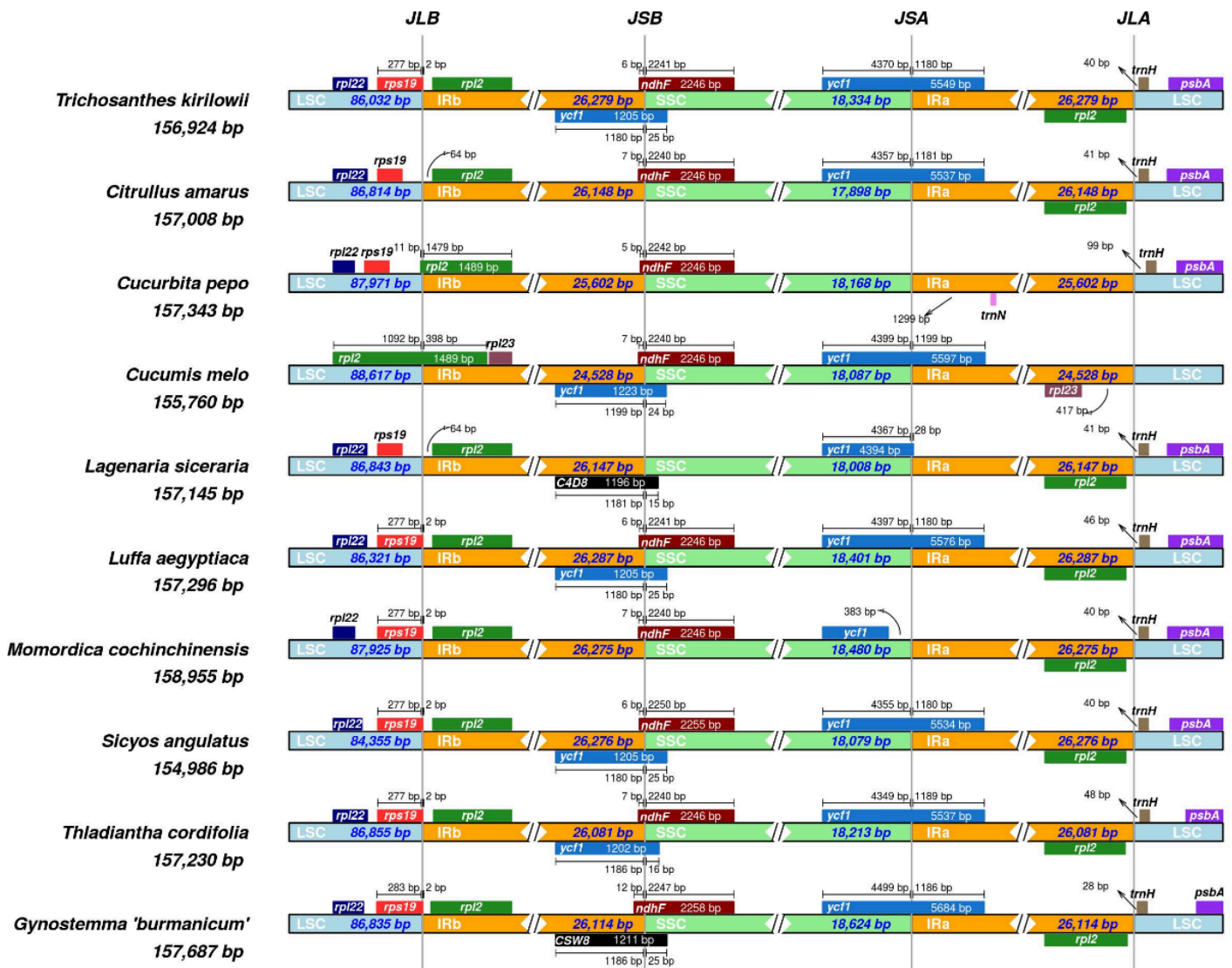


Figure 4

Comparison of the junctions among the LSC, SSC, IRa, and IRb regions across ten chloroplast genomes in the Cucurbitaceae family.

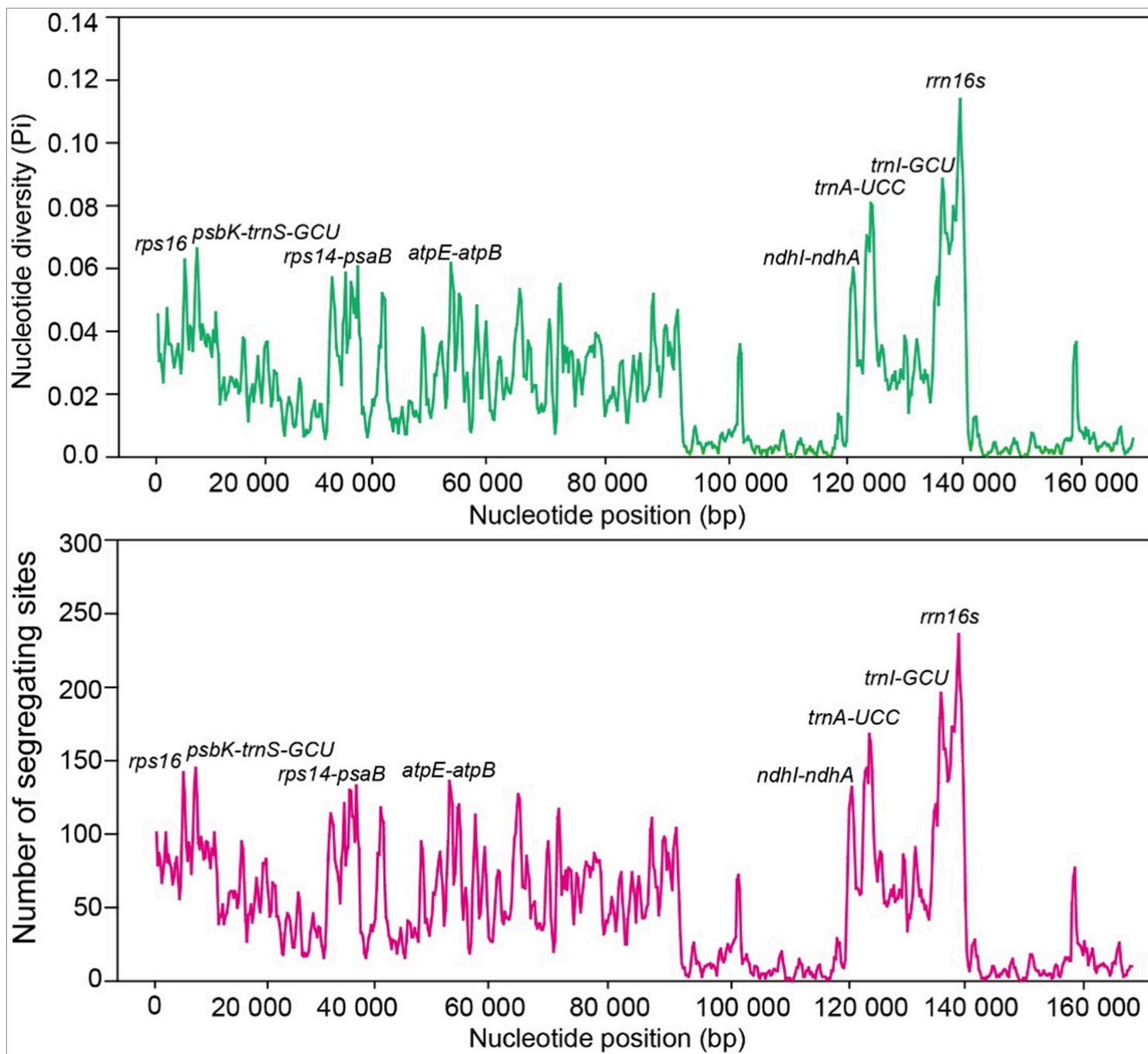


Figure 5

Sliding window analysis of the complete cp. genomes of Cucurbitaceae family

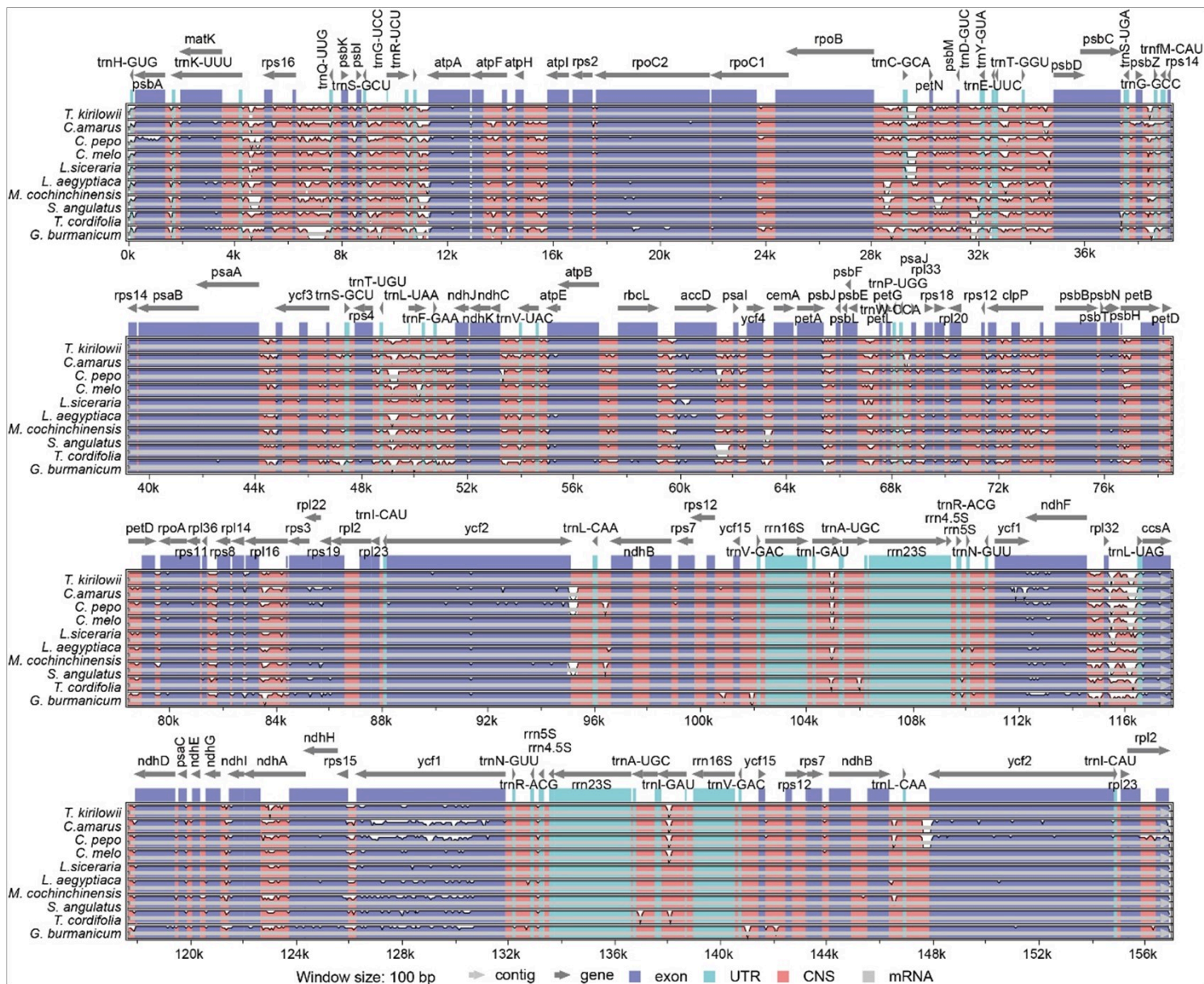


Figure 6

Sequence alignment plot for 10 Cucurbitaceae species using mVISTA with *T. kirilowii* chloroplast genome as a reference.

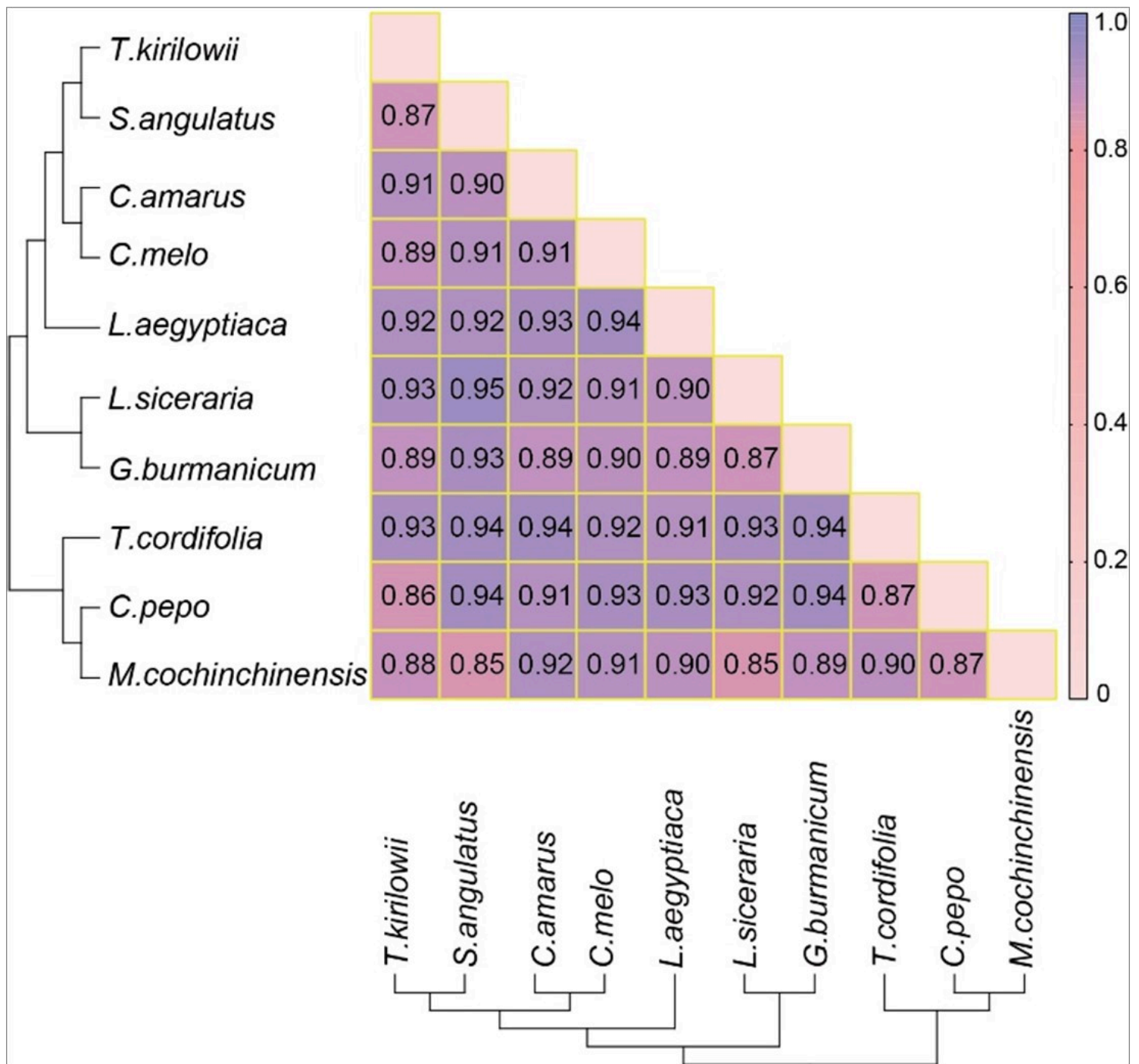


Figure 7

The Ka/Ks values of shared protein-coding genes across the chloroplast genomes of 10 Cucurbitaceae species

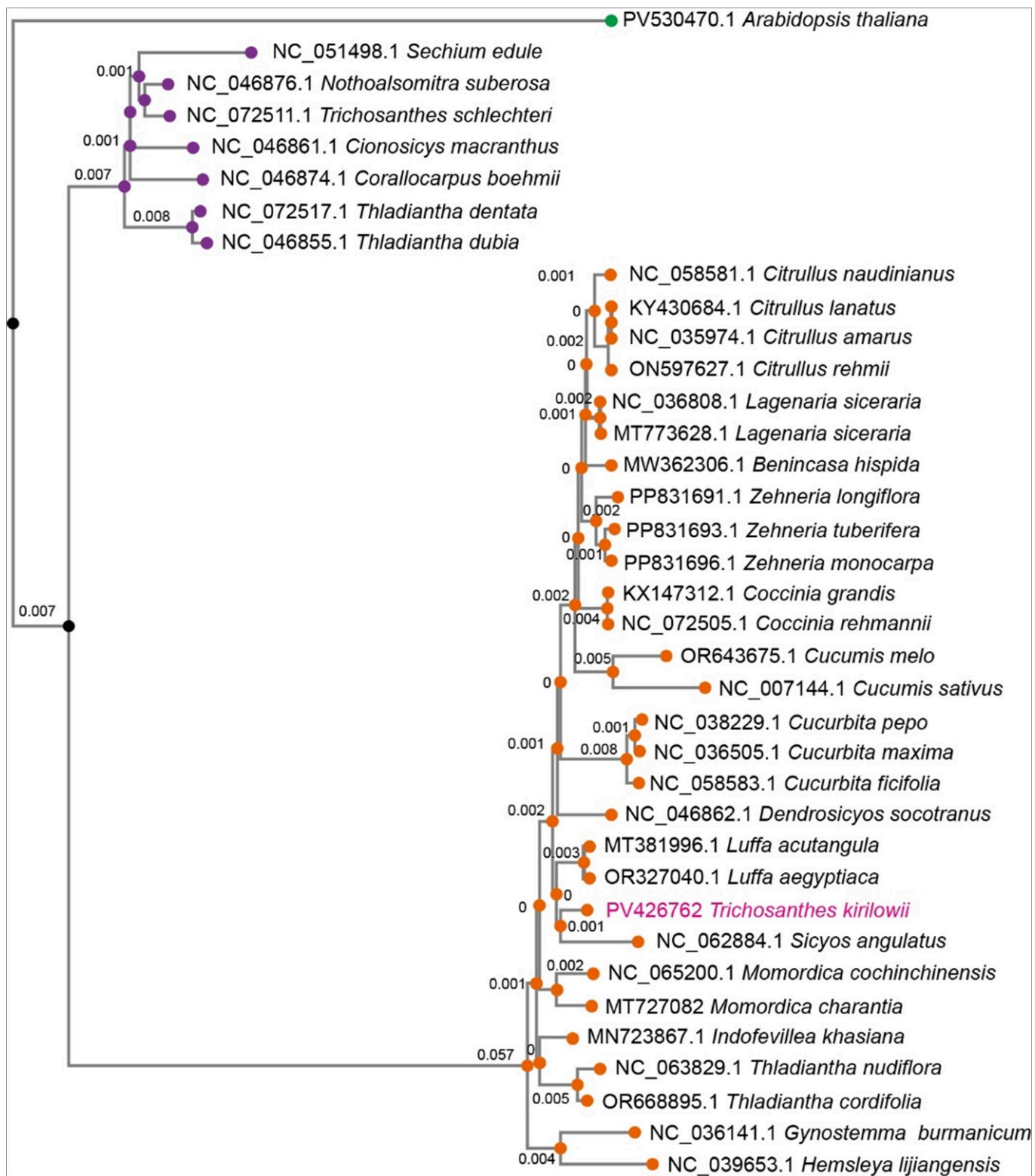


Figure 8

The evolutionary phylogeny of 36 Cucurbiaceae species was constructed based on chloroplast genome sequences, with *Arabidopsis thaliana* serving as the outgroup.