Genomics Communications

ARTICLE

Open Access

https://doi.org/10.48130/gcomm-0024-0005

Genomics Communications **2024**, 1: e005

Long repeat sequences mediated multiple mitogenome conformations of mulberries (*Morus* spp.), an important economic plant in China

Luxian Liu[1,2#*], Qian Long[1,2#], Weiwei Lv[2], Jiayi Qian[3], Ashley N. Egan[4], Yu Shi[2] and Pan Li[5*]

[1] *College of Life Sciences, Henan Normal University, Xinxiang 453000, China*
[2] *Laboratory of Plant Germplasm and Genetic Engineering, School of Life Sciences, Henan University, Kaifeng 475001, China*
[3] *College of Life Sciences, Henan Agricultural University, Zhengzhou 450002, China*
[4] *Department of Biology, Utah Valley University, Orem, UT 84058, USA*
[5] *Systematic & Evolutionary Botany and Biodiversity Group, MOE Key Laboratory of Biosystems Homeostasis and Protection, College of Life Sciences, Zhejiang University, Hangzhou 310058, China*

[#] Authors contributed equally: Luxian Liu, Qian Long

* Corresponding authors, E-mail: liushuangcx2007@126.com; panli_zju@126.com

Abstract

Mulberries (genus *Morus*; Moraceae) hold significant economic value in sericulture and have great potential in the horticultural industry, food industry, and human health arenas worldwide. Since the advent of the genomic era, biological macromolecules of *Morus* species such as whole genome and plastome sequences have been reported, but mitochondrial genome sequences are relatively scarce which greatly hinders the comprehensive understanding of the evolutionary history and processes at work with *Morus*. Here, four *Morus* mitogenomes were assembled using Illumina and PacBio HiFi data. The results elucidated that the structure of the four mitogenomes was greatly heterogeneous due to the presence of different numbers of repeat-mediated recombination events with multiple conformations existing simultaneously in the mitogenome for each species. The genome size ranged from 359,062 to 376,846 bp. The repeat sequences and gene transfers from plastome to mitogenome varied widely among the four mitogenomes, which was the main cause of variation in mitogenome size. Finally, the evolutionary history of Rosales was inferred based on the mitogenome sequences. The analyses revealed a strong difference in the phylogenetic placement of Rhamnaceae compared to earlier plastid or nuclear phylogenies, likely due to the effects of ancient hybridizations. Overall, the results presented here will provide important genetic resources for the utilization of this important economic plant.

Citation: Liu L, Long Q, Lv W, Qian J, Egan AN, et al. 2024. Long repeat sequences mediated multiple mitogenome conformations of mulberries (*Morus* spp.), an important economic plant in China. *Genomics Communications* 1: e005 https://doi.org/10.48130/gcomm-0024-0005

Introduction

Morus L., commonly known as mulberry trees, is a genus of flowering plants in the family Moraceae. These deciduous trees are widely distributed and/or naturalized in all temperate areas across the globe, also in the mountains of Indonesia and South America[1]. Mulberry trees typically have broad, lobed leaves and produce small, sweet, multiple-fruit aggregates that resemble berries, with the fruits varying in color depending on the species, ranging from white to deep purple-black[2]. As the sole food consumed by the silkworm (*Bombyx mori*), mulberry represents a vital component of sericulture, which originated more than 5,000 years ago and strongly influenced world history and the expansion of a global economy[3–5]. Mulberry holds potential beyond its use of foliage to feed the silkworm, particularly in the horticultural industry, food industry, and human health applications worldwide[6,7]. Various parts of mulberry trees, including leaves, barks, and fruits, have been utilized in traditional medicine for antioxidant, anti-inflammatory, and antimicrobial properties[8,9]. Beyond medicine, mulberry trees offer lush foliage, colorful seasonal changes, delicious fruit, wildlife attraction, adaptability to diverse environments, graceful form, and rich historical significance, making them a desirable addition to any native landscape[2,10].

Since Linnaeus first established the genus *Morus* with the description of seven species[11], over 150 *Morus* species names have been cited in the Index Kewensis to date[12]. However, 17 of these species are currently accepted by the vast majority of botanical authorities[1]. Interspecies hybridization occurs frequently in *Morus*[4], a phenomenon reported many times in previous studies. Hybridization has been detected between white mulberry (*M. alba*) and red mulberry (*M. rubra*), and asymmetrical introgression of these two species directly led to the decline of the local populations of red mulberry[13,14]. Hybridizations have also been documented between *M. alba* and *M. australis*[15], *M. alba* and *M. serrata*[16], as well as between *M. boninensis* and *M. australis*[17]. Frequent hybridization, combined with a wide geographical distribution, a long domestication history, and considerable morphological plasticity, make the taxonomy of *Morus* challenging[5,18]. With the increasing amounts of genomic data and many newly developed bioinformatic pipelines, more and more low or single-copy nuclear genes and plastomes are available for species identification and phylogeny of different plant lineages[19–22]. Our scientific understanding of *Morus* has benefited from these innovations, with the phylogeny and taxonomy of *Morus* successfully resolved[5] based on nuclear and plastid data.

Mitochondria play a vital role in the growth, development, programmed cell death, and male sterility of angiosperms[23,24]. Unlike most plastomes with stable size and gene content, the plant mitochondrial genome (mitogenome) is extremely variable in its size, collinearity, number of repeats, and general structure[25]. The size of the plant mitogenome ranges from 66 kb[26] to nearly 12 Mb[27] across different species, and the mitogenomes exhibit circular, linear, and even complex branching and reticular structures[28,29]. Most changes in the size and structure of mtDNA are thought to be caused by repeat-mediated duplication and recombination[30], and frequent gene transfer between organelle genomes and between organelle and nuclear genomes aggravate these changes[31,32]. Consequently, the assembly and conformational determination of the mitogenome can be more complex and challenging than that of plastomes.

Innovations with the genomics era has brought new opportunities for mitogenome research and promoted the rapid development of mitochondrial genetics and genomics. Different pipelines have been developed to assemble the mitogenome, such as GetOrganelle[33], GSAT[34], and PMAT[35], which facilitated the generation of reliable plant mitogenome assembly graphs. A series of studies have indicated that mitogenomes can also be used for population genetics, molecular ecology, plant phylogeny, and adaptive evolution[36–39]. To date, plastomes from 4,019 genera and 16,435 species have been published based on the Chloroplast Genome Information Resource website (https://ngdc.cncb.ac.cn/cgir/statistic), but only 673 plant mitogenomes are reported according to the result of Wang et al.[40]. For the genus *Morus*, a total of 104 plastomes from 13 species have been assembled. Still, data on the mitogenomes are lacking with only three incomplete mitogenomes involving two species (*M. alba* and *M. notabilis*) released in GenBank as of June 2024. This limitation significantly hinders research on the evolution of *Morus* mitogenomes.

The current study sequenced and assembled the complete mitogenomes of three *Morus* species: *M. alba*, *M. mongolica*, and *M. notabilis*. Using Illumina short reads and PacBio HiFi long reads, the dynamic transformation of mitogenome conformation was investigated across these species. Their genomic characteristics, including gene content, intracellular gene transfer, and organellar phylogeny were also compared. These findings enhance our understanding of the evolution of organelle genomes in *Morus* species and offer insights into their phylogeny, germplasm resource identification, genetic diversity conservation, and environmental adaptation mechanisms.

Materials and methods

Plant materials and sequencing

Fresh leaves from one individual of *Morus mongolica* were collected from the Baligou scenic area, Henan, China (113°34′25.12″ E, 35°36′26.66″ N), immediately frozen in liquid nitrogen and then stored at −80 °C. Dr. Luxian Liu undertook the identification of *M. mongolica*, and a voucher specimen was deposited at the Herbarium of Henan University with the accession number LVV20240608. No specific permissions were required for sample collection which are neither privately owned nor protected and the field study did not involve endangered or protected species. Genomic DNA was extracted using a QIAGEN Genomic Kit (Qiagen, CA, USA). Illumina paired-end (PE) reads were generated on the Illumina NovaSeq 6000 (Illumina Inc., San Diego, CA, USA) and long-read sequencing was conducted using the PacBio Revio sequencer (Pacific Biosciences of California, Inc., USA). All of the above sequencing protocols were performed by the Genome Center of NextOmics (Wuhan, China). To further explore the mitogenome evolution of *Morus*, HiFi data and Illumina data of two *M. alba* cultivars and one *M. notabilis* were downloaded from the National Genomics Data Center (Supplementary Table S1).

Assembly and annotation of organelle genomes

For the plastome assembly, GetOrganelle v1.7.5.1[33] was used to assemble the Illumina paired-end reads with default parameters directly. For the mitogenome assembly, the 'autoMito' module in PMAT v1.5.1[35] software was utilized to conduct *de novo* assembly of the PacBio HiFi data derived from *Morus* species. Next, 24 conserved plant mitochondrial protein-coding genes (PCGs) from the 15 representative species were used to identify the draft mitogenome by the BLASTn v2.2.30+[41] based on the assembled contigs. Then, the mitogenome graph was simplified using Bandage v 0.8.1[42] and the plastid or nuclear contigs removed. The Illumina paired-end reads were mapped to the mitogenome using CLC Genomics Workbench 12 (CLC Inc. Aarhus, Denmark) to improve assembly accuracy.

The plastomes of *Morus* were annotated in Geneious R11[43], and putative starts stops, and intron positions of genes were corrected by comparison with homologous genes of the *M. alba* plastome (GenBank accession number: OQ644502). The mitogenomes were annotated with the IPMGA webserver (www.1kmpg.cn/mga), and all annotations were further manually checked and corrected for start and stop codons of PCGs using Geneious R11. The tRNA genes were identified using tRNAscan-SE v1.21[44]. Finally, the plastome and mitogenome maps were generated using the OrganellarGenomeDRAW tool (OGDRAW v1.3.1)[45] followed by manual modification.

Repeat-mediated homologous recombination prediction

In the initial assembly, there were multiple repeats present in the mitogenomes, which resulted in different mitogenome conformations for *Morus* species. Therefore, the following strategy for each repeat was used to validate the potential conformations of the mitogenome. First the sequences of the repetitive elements and their flanking regions of 1 kb from the genome assembled by PMAT were extracted to ensure that long reads can completely cross the repetition. Subsequently, the long reads were mapped to the above extracted sequences using BWA v0.7.17[46], and the alignment file was converted to BAM format using SAMtools v1.13[47]. Finally, each long read was carefully checked using Tablet[48] and the number of long reads that completely crossed the repeat region counted. The conformation with the largest number of long reads was defined as the dominant conformation. For the branch nodes, a 2 kb region on each side of the node was extracted as the reference sequence and the BAM file obtained was visualized using the above method by IGV v2.17.4[49]. The depth at the node was calculated and the path with higher depth selected as the dominant conformation.

Analysis of RSCU and repeated sequences

The PCGs of the mitogenome were extracted using PhyloSuite v1.2.3[50], and the relative synonymous codon usage (RSCU) in protein coding sequences of *Morus* was investigated using the program CodonW v 1.4.2 (http://codonw.sourceforge.net/).

Forward, reverse, palindromic, and complementary repeats were detected using the online REPuter software (https://bibiserv.cebitec.uni-bielefeld.de/reputer)[51] with maximum computed repeats, minimal repeat size, and hamming distance set to 5,000, 30, and 3 respectively. The tandem repeats with longer than 6 bp repeat units were identified using Tandem Repeats Finder v.4.09 (http://tandem.bu.edu/trf/trf.submit.options.html)[52] with default parameters. Simple sequence repeats (SSRs) were detected using MISA (https://webblast.ipk-gatersleben.de/misa/)[53], and a threshold of ten, five, four, three, three, and three repeat units for mono-, di-, tri-, tetra-, penta-, and hexanucleotide SSRs, were applied respectively.

Identification of mitochondrial plastid DNA (MTPT)

To identify plastid-derived DNA fragments in the mitogenome, the plastome and mitogenome sequences were compared for each *Morus* species using BLASTn[41] with the following parameters: -evalue 1e-5, -word_size 9, -gapopen 5, -gapextend 2, -reward 2, -penalty -3, and -dust no. The BLASTn results were visualized using the Circos package[54], and the identified MTPTs were also annotated using the IPMGA webserver.

Phylogenetic inference of Rosales

In phylogeny inference, 40 Rosales accessions including four *Morus* accessions analyzed in this study, and another 36 Rosales species

downloaded from GenBank were selected to reconstruct phylogenetic trees with two Cucurbitales species selected as the outgroups (Supplementary Table S2). To minimize the impact of missing data or insufficiently informative loci on phylogenetic inference, the minimum number of species that share PCGs was set to 35. Maximum likelihood (ML) and Bayesian inference (BI) were performed to estimate the phylogeny based on all shared PCGs. PARTITIONFINDER v2.1.1[55] was used to determine the optimal nucleotide substitution models. ML analysis was performed using RAxML v8.1.11[56] with 1000 rapid bootstrap iterations and the default settings for other parameters. BI analysis was carried out using MrBayes v3.2.3[57], and the posterior probability was estimated using four chains running 5,000,000 generations sampling every 1000 generations. Convergence of the MCMC chains was assumed when the average standard deviation of split frequencies reached 0.01 or less, and the first 25% of sampled trees were considered burn - in trees.

Results

Characteristics of the mitogenomes of *Morus* species

The mitogenomes of *Morus* species showed a complex assembly graph mediated by different numbers of long repeats (LRs) or node branches based on long-reads data (Fig. 1). One pair of LRs existed in each of the mitogenome of *M. notabilis* and *M. alba* 'Zhongsang5801' (*M. alba*-ZS5801; Fig. 1a, c), three pairs of LRs occurred in *M. alba* 'Zhenzhubai' (*M. alba*-ZZB; Fig. 1d), and LRs-mediated multiple conformations were not detected in *M. mongolica* (Fig. 1b). These repeats were resolved by artificially simulating four possible paths and making judgments based on the mapping results of long reads, and the dominant conformation of the mitogenome for *Morus* species was obtained by using this strategy. For *M. notabilis*, the repeat sequence (LR1) with 9,951 bp length mediated genomic recombination and formed one master circular structure or two small circular structures with equal probability (Fig. 2a), and here the master circle representing the complete

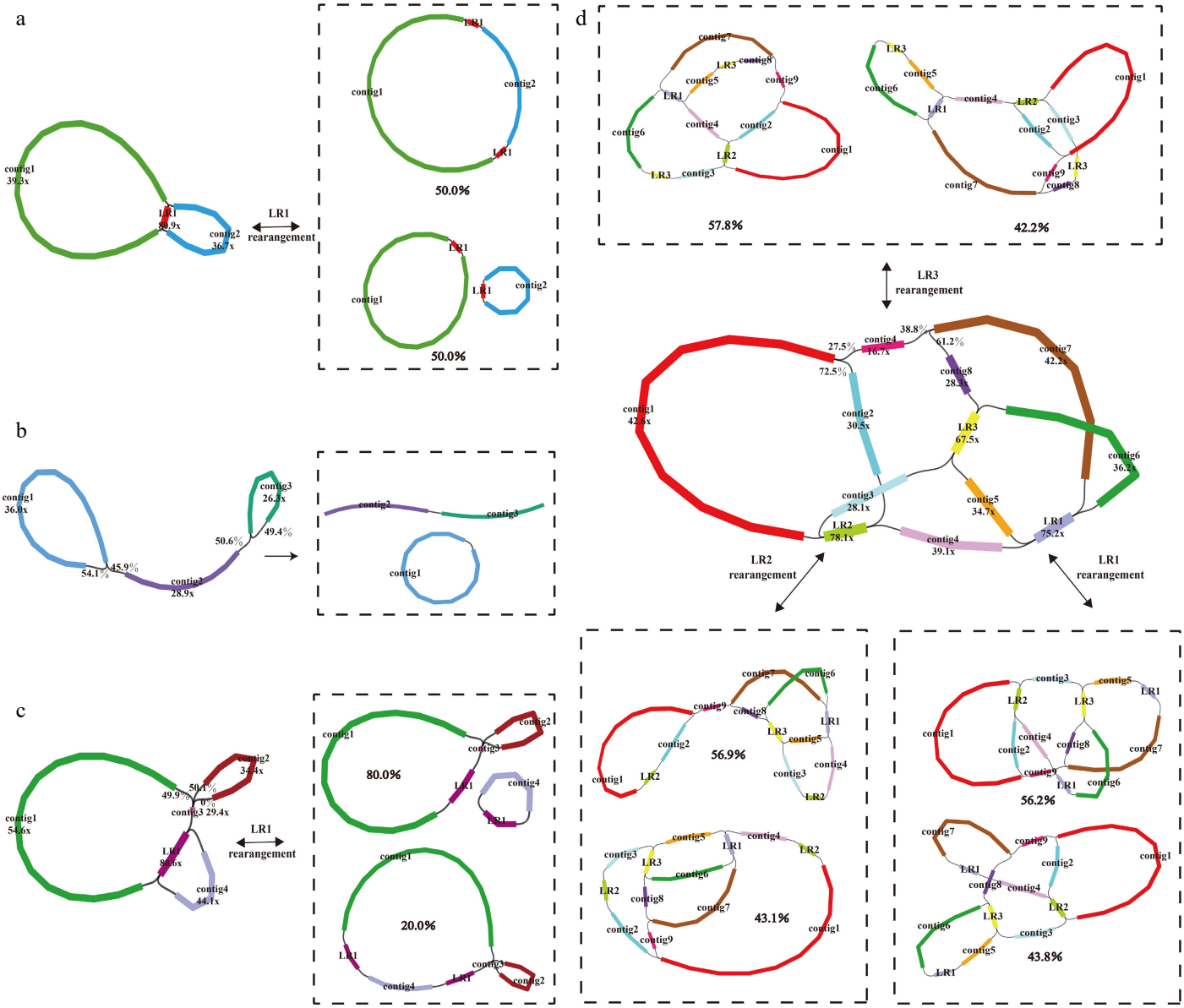

Fig. 1 Mitogenome assembly graph and possible connections (black lines) mediated by repeats for (a) *M. notabilis*, (b) *M. mongolica*, (c) *M. alba*-ZS5801, and (d) *M. alba*-ZZB. Each colored segment is labeled with its name and coverage. The boxes represent the different mitogenome conformations mediated by each pair of repeats and the probability of being supported.

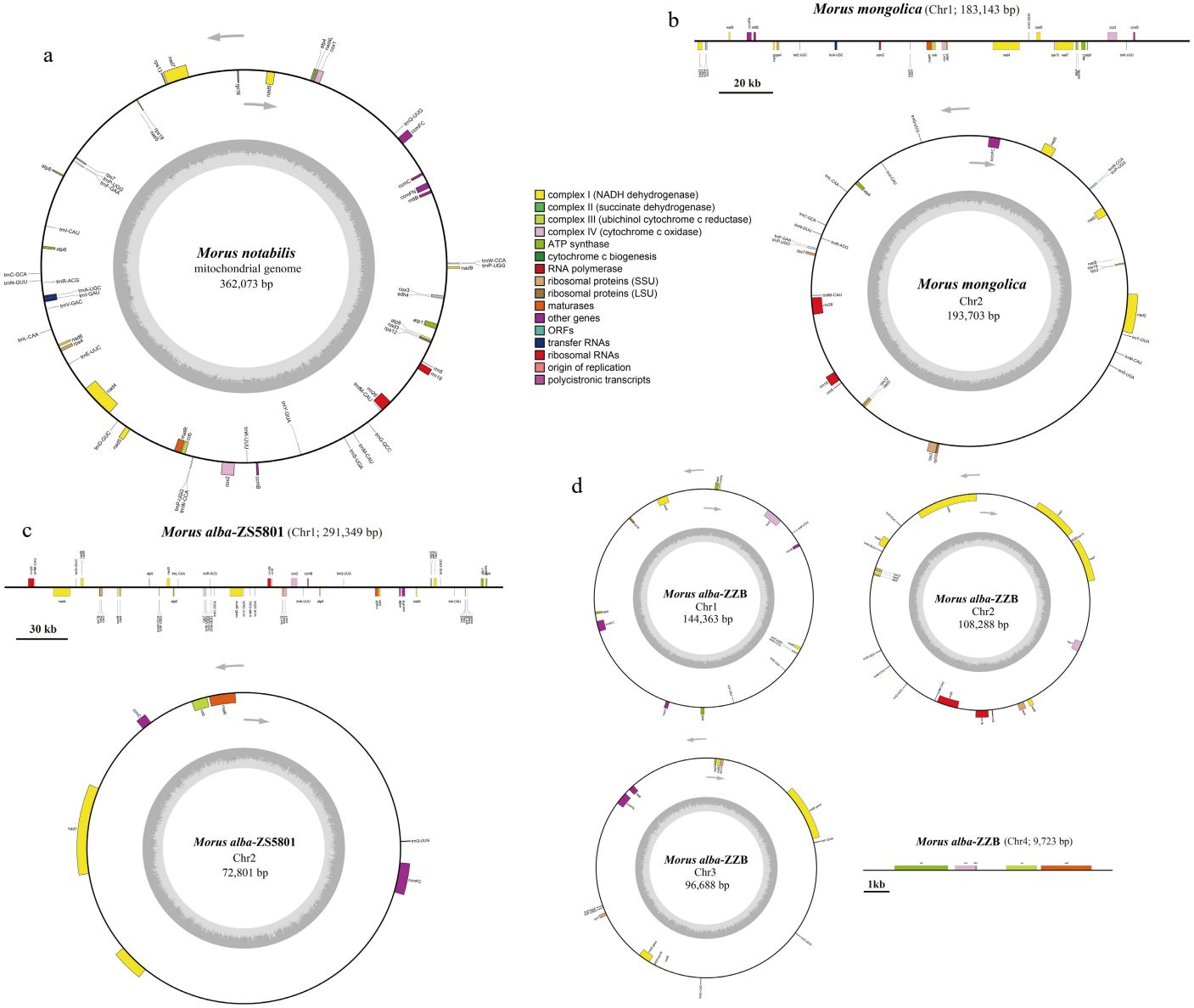

Fig. 2 The mitogenome map of *Morus* species. (a) *M. notabilis,* (b) *M. mongolica,* (c) *M. alba*-ZS5801, and (d) *M. alba*-ZZB. Genes belonging to different functional groups are color-coded.

mitogenome was used for subsequent analysis. In the mitogenome of *M. mongolica*, the obtained complex structure was disassembled into a linear molecule and a circular molecule after calculating the depth at the node (Fig. 2b). The two cultivars of *M. alba* showed two completely different mitogenomic conformations, the genomic recombination mediated by one pair of LRs (LR1) resulted in the mitogenome disassembled into a linear molecule and a circular molecule in the cultivar 'Zhongsang5801' (Fig. 2c), and three pairs of LRs (LR1, LR2, and LR3) occurred in the mitogenome of the cultivar 'Zhenzhubai' which divided the complex structure into three circular molecules and one linear molecule (Fig. 2d). Finally, the paired-end reads and long reads were mapped to the mitogenomes, and the statistics of the sequencing depth showed that a gap-free genome assembly of high quality was obtained (Supplementary Fig. S1).

The complete mitogenomes of the four *Morus* accessions ranged from 359,062 bp (*M. alba*-ZZB) to 376,846 bp (*M. mongolica*), revealing a large interspecific variation among different *Morus* species and cultivars (Fig. 2). The GC content for the four mitogenomes varied slightly from 45.4% (*M. alba*-ZS5801) to 45.7% (*M. notabilis*). The four *Morus* mitogenomes encoded a different set of genes with the number ranging from 49 to 55, including 31–32 PCGs, 15–20 tRNA genes and three rRNA genes (Supplemental Fig. S2, Table 1). Among these genes, 30 PCGs, 14 tRNA genes, and all three rRNA genes (*rrn*5, *rrn*18, and *rrn*26) were shared by all four mitogenomes. Seven PCGs including *ccm*FC, *cox*2, *nad*1, *nad*2, *nad*4, *nad*5, *nad*7, and two tRNAs (*trn*A-UGC, *trn*I-GAU) contained more than one intron (Table 1). For PCGs, *rps*3 and *rps*13 were lost in *M. alba*-ZZB and *M. alba*-ZS5801 respectively, three PCGs including *atp*9, *nad*3, and *rps*12 had two copies in *M. alba*-ZZB and two genes *cob* and *mat*R had also two copies in *M. alba*-ZS5801. For tRNA genes, three, five, and four tRNA genes were missing in *M. mongolica*, *M. alba*-ZZB, and *M. alba*-ZS5801 respectively, only *trn*Q-UUG had two copies in *M. alba*-ZS5801 and two tRNA genes (*trn*P-UGG, *trn*W-CCA) had one more copy in *M. notabilis* compared with the other three genomes.

Repeat sequences analysis

The dispersed repeats within each *Morus* mitogenome were identified as two major types, including 105 (*M. notabilis*)–141 (*M. mongolica*)

Table 1. Summary of genes contained in the four *Morus* mitogenomes.

Group of genes		*M. notabilis*	*M. mongolica*	*M. alba-ZS5801*	*M. alba-ZZB*
Core genes	ATP synthase	*atp*1, *atp*4, *atp*6, *atp*8, *atp*9	*atp*1, *atp*4, *atp*6, *atp*8, *atp*9	*atp*1, *atp*4, *atp*6, *atp*8, *atp*9	*atp*1, *atp*4, *atp*6, *atp*8, *atp*9 (2)
	Cytochrome c biogenesis	*ccm*B, *ccm*C, *ccm*FC*, *ccm*FN	*ccm*B, *ccm*C, *ccm*FC*, *ccm*FN	*ccm*B, *ccm*C, *ccm*FC*, *ccm*FN	*ccm*B, *ccm*C, *ccm*FC*, *ccm*FN
	Ubichinol cytochrome c reductase	*cob*	*cob*	*cob* (2)	*cob*
	Cytochrome c oxidase	*cox*1, *cox*2*, *cox*3	*cox*1, *cox*2*, *cox*3	*cox*1, *cox*2*, *cox*3	*cox*1, *cox*2*, *cox*3
	Maturases	*mat*R	*mat*R	*mat*R (2)	*mat*R
	Transport membrane protein	*mtt*B	*mtt*B	*mtt*B	*mtt*B
	NADH dehydrogenase	*nad*1*, *nad*2****, *nad*3, *nad*4***, *nad*4L, *nad*5**, *nad*6, *nad*7****, *nad*9	*nad*1*, *nad*2***, *nad*3, *nad*4***, *nad*4L, *nad*5**, *nad*6, *nad*7****, *nad*9	*nad*1*, *nad*2***, *nad*3, *nad*4***, *nad*4L, *nad*5**, *nad*6, *nad*7****, *nad*9	*nad*1*, *nad*2***, *nad*3 (2), *nad*4***, *nad*4L, *nad*5**, *nad*6, *nad*7****, *nad*9
Variable genes	Large subunit of ribosome	*rpl*16	*rpl*16	*rpl*16	*rpl*16
	Small subunit of ribosome	*rps*3*, *rps*4, *rps*7, *rps*12, *rps*13, *rps*19	*rps*3, *rps*4, *rps*7, *rps*12, *rps*13, *rps*19	*rps*3, *rps*4, *rps*7, *rps*12, *rps*19	*rps*4, *rps*7, *rps*12 (2), *rps*13, *rps*19
	Succinate dehydrogenase	*sdh*4	*sdh*4	*sdh*4	*sdh*4
rRNA genes	Ribosomal RNAs	*rrn*5, *rrn*18, *rrn*26	*rrn*5, *rrn*18, *rrn*26	*rrn*5, *rrn*18, *rrn*26	*rrn*5, *rrn*18, *rrn*26
tRNA genes	Transfer RNAs	*trn*A-UGC*, *trn*C-GCA, *trn*D-GUC, *trn*E-UUC, *trn*F-GAA, *trn*fM-CAU, *trn*G-GCC, *trn*I-CAU, *trn*I-GAU*, *trn*K-UUU, *trn*L-CAA, *trn*M-CAU, *trn*N-GUU, *trn*P-UGG (3), *trn*Q-UUG, *trn*R-ACG, *trn*S-UGA, *trn*V-GAC, *trn*W-CCA (2), *trn*Y-GUA	*trn*A-UGC*, *trn*C-GCA, *trn*D-GUC, *trn*E-UUC, *trn*F-GAA, *trn*fM-CAU, *trn*I-CAU, *trn*K-UUU, *trn*L-CAA, *trn*M-CAU, *trn*N-GUU, *trn*P-UGG (2), *trn*Q-UUG, *trn*R-ACG, *trn*S-UGA, *trn*W-CCA, *trn*Y-GUA	*trn*C-GCA, *trn*D-GUC, *trn*E-UUC, *trn*F-GAA, *trn*fM-CAU, *trn*I-CAU, *trn*K-UUU, *trn*L-CAA, *trn*M-CAU, *trn*N-GUU, *trn*P-UGG (2), *trn*Q-UUG (2), *trn*R-ACG, *trn*S-UGA, *trn*W-CCA, *trn*Y-GUA	*trn*C-GCA, *trn*D-GUC, *trn*E-UUC, *trn*F-GAA, *trn*fM-CAU, *trn*G-GCC, *trn*I-CAU, *trn*K-UUU, *trn*M-CAU, *trn*N-GUU, *trn*P-UGG (2), *trn*Q-UUG, *trn*S-UGA, *trn*W-CCA, *trn*Y-GUA

The number of asterisks indicate the number of introns that the genes contained. The number in the parentheses indicate the number of copies of the genes.

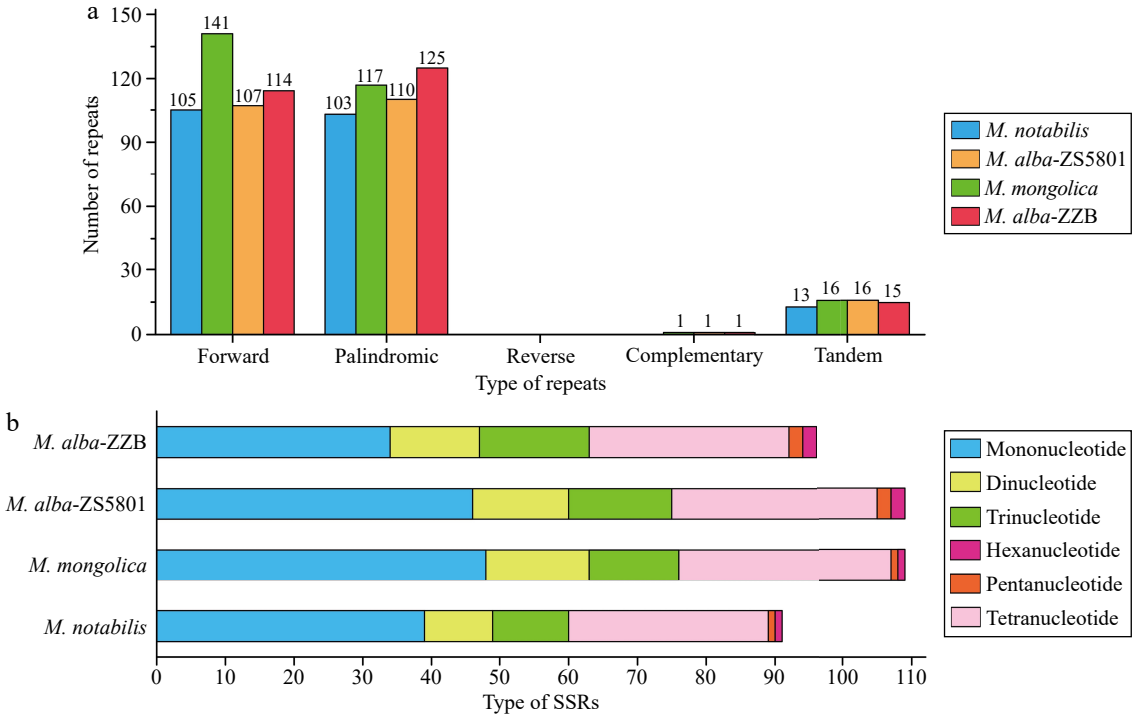

Fig. 3 Repeat sequences in the *Morus* mitogenomes. (a) Type and number of dispersed repeats and tandem repeats, (b) type and number of simple sequence repeats (SSRs).

forward and 103 (*M. notabilis*)–125 (*M. alba*-ZZB) palindromic repeats and one complementary repeat was found in three mitogenomes except *M. notabilis* (Fig. 3a). No reverse repeat sequences were observed in any of the four mitogenomes. These repeats ranged from 30 to 22,006 bp (eight were longer than 1 kb) (Supplementary Table S3), with most distributed in non-coding regions and some found in PCGs such as *ccmFC*, *nad2*, *nad4*, *nad5*, *nad7*, *cob*, *cox2*, *matR*, and *atp9*. Additionally, 13–16 tandem repeats with repeat lengths ranging from 12–39 bp were detected in the four genomes and most of these tandem repeats had a length of 15–25 bp (Fig. 3a; Supplementary Table S4). Moreover, 91 (*M. notabilis*)-109 (*M. alba*-ZS5801) microsatellites were detected across the four *Morus* mitogenomes, comprising of 34–48 mono-, 10–15 di-, 11–16 tri-, 29–31 tetra-, 1–2 penta-, and 1–2 hexanucleotide repeat motifs (Fig. 3b).

Codon bias analysis of mitogenome

Codon usage was analyzed in the 32 PCGs of the *M. notabilis* and *M. mongolica* mitogenomes, 33 PCGs of *M. alba*-ZS5801, and 34 PCGs of *M. alba*-ZZB. A total of 9,619, 9,629, 10,552, and 9,617 codons were identified and a high similarity in codon usage and amino acid frequencies was observed in all four mitogenomes (Fig. 4; Supplementary Table S5). Most of the PCGs used AUG as the start codon, while GUG and AUA were found to be the start codon for *rpl*16 and *nad*1 respectively in all four accessions. Three stop codons including UAA, UAG, and UGA were identified in the PCGs, but UAA with RSCU > 1 was the most preferred codon. For the codons of encoding amino acids, there were 31 (*M. notabilis*), 30 (*M. mongolica*), 33 (*M. alba*-ZS5801), and 31 (*M. alba*-ZZB) codons with RSCU > 1, indicating a higher frequency of these codons compared to other synonymous codons. Among them, the GCU codon for Ala had the highest RSCU value (1.58−1.64), as well as the CAG codon for Gln had the smallest one with an RSCU value ranging from 0.44 to 0.47.

Plastid DNA transfer to mitogenome (MTPTs)

The plastomes of the four *Morus* accessions assembled in this study varied slightly in length, ranging from 159,136 bp (*M. mongolica*) to 159,546 bp (*M. notabilis*; Fig. 5). To eliminate redundant detections, only a single IR region of each plastome was used for analysis, and 38, 39, 35, and 40 mitochondrial-plastid DNA transfers (MTPTs) were identified in the mitogenomes of *M. notabilis*, *M. mongolica*, *M. alba*-ZS5801, and *M. alba*-ZZB, respectively (Fig. 5; Supplementary Table S6). These MTPTs had a combined length of 26,024 bp (*M. notabilis*), 21,398 bp (*M. mongolica*), 17,698 bp (*M. alba*-ZS5801), and 10,112 bp (*M. alba*-ZZB) accounting for 7.19%, 5.68%, 4.86%, and 2.82% of the mitogenomes. Among these MTPTs, the majority of transferred fragments for each accession were shorter than 1000 bp and there was a total of 14 MTPTs longer than 1,000 bp in all the mitogenomes, both the longest (16,238 bp) and shortest (30 bp) MTPTs were found in *M. notabilis*. These MTPTs were then extracted and annotated. Five intact tRNA genes were discovered (*trn*P-UGG, *trn*W-CCA, *trn*D-GUC, *trn*M-CAU, and *trn*N-GUU) and were identified as being transferred from plastomes to the mitogenomes for all accessions. Two intact tRNA genes including *trn*L-CAA and *trn*R-ACG transferred to three of the four mitogenomes except *M. alba*-ZZB, *trn*I-GAU, and *trn*V-GAC were only transferred in *M. notabilis* as well as *trn*A-UGC was transferred in *M. notabilis* and *M. mongolica*. In addition, *trn*F-GAA and *trn*Q-UUG might have experienced rapid sequence divergences during the migration process due to partial gene fragments identified in the four mitogenomes. For PCGs, only some fragments of *rps*12 were transferred to the four mitogenomes.

Phylogenetic analyses

For the phylogenetic inference, 25 individual gene alignments of PCGs were concatenated to generate a combined matrix comprising 42 accessions with a length of 24,627 bp. The best-fit nucleotide substitution model determined for the concatenated matrix was GTR + I + G. The tree topologies reconstructed by ML and BI analyses were completely consistent (Fig. 6). In the phylogenetic trees, most nodes were strongly supported with maximal ML bootstrap (BS) or Bayesian posterior probability (PP), and all the species were grouped into six families with different species of the same genus forming monophyletic clades. In the Rosales, the six families were divided into two clades with maximal support (BS/PP = 100/1). Two subfamilies of Rosaceae, Amygdaloideae, and Rosoideae, comprised one clade. In the other clade (BS/PP = 100/1), the family Rhamnaceae was the first diverging lineage, followed by a clade comprising two *Hippophae* species from the family Elaeagnaceae. For the remaining three families, Moraceae and Cannabaceae first formed a monophyletic group (BS/PP = 100/1) which was subsequently sister to the family Ulmaceae. Within the family Moraceae, ten species from five genera grouped into two clades. Four individuals of three *Morus* species formed one of the clades (BS/PP = 100/1), *M. notabilis* was firstly diverged and two individuals of *M. alba* were not resolved as monophyletic. *Ficus carica* occupied the basal position of the other clade with moderate support (BS/PP = −/0.76), followed by a maximal supported clade (BS/PP = 100/1) comprising of *Malaisia scandens* and *Allaeanthus kurzii* sister then to the genus *Broussonetia*. For the three *Broussonetia* species, *B. papyrifera* was well-supported to be sister to the *B. kaempferi* and *B. monoica* clade.

Discussion

Complex mitogenome structure in *Morus* species

Compared to the mitogenome of animals and the plastome of plants, the plant mitogenome underwent dramatic structural changes during evolution, which involve huge genome size variations, genomic rearrangements, repeat-mediated recombination, low gene density, and gene transfer[58,59]. Due to the presence of repeat-mediated recombination, plant mitogenomes usually exhibit multiple conformations[59,60], including circular (e.g. *Crataegus* spp.)[24], linear and branched (e.g. *Lactuca* spp.)[29], and numerous smaller circular molecules (e.g. *Actinidia chinensis*)[61]. Previous studies have reported that the variations in mitogenomes are not only between plant species but also can be within the same species[62,63], and even plant mtDNA may simultaneously exist in different genome conformations within the same individual[64]. Given the importance of the plant mitogenome in understanding the phylogeny, germplasm resource identification, genetic diversity conservation and adaptive evolution, four representative accessions of three *Morus* species were sequenced and comparatively analyzed in this study.

The complete mitogenome sequences of the four *Morus* accessions ranged from 359,062 to 376,846 bp in length, and encoded a different set of genes with the number ranging from 49 to 55 (Fig. 2). The size and gene content of mitogenomes within *Morus* varied more than that of plastomes sequenced in this study and compared with those published by a previous study[5]. Upon further confirmation, it was found that the variation of mitogenome size of *Morus* species was mainly caused by the impact of repetitive sequences within it (Fig. 3), rather than foreign sequences transferred from the plastomes (Fig. 5). Compared with the reported mitogenomes of other Moraceae species, the relatively medium size of *Morus* was larger than that of *Broussonetia* (276,967–325,822 bp)[65] but smaller than *Ficus* (480,902 bp)[66].

Previous studies have shown that the plant mitogenome has a dynamic structure due to recombination events that are commonly facilitated by repetitive sequences dispersed across the mitochondrial DNA[67–69]. In the assembly of the *Morus* mitogenomes, repeat-mediated recombination activity was also found. One or three repeat sequences were confirmed that mediated recombination and generated multiple genome conformations in *Morus* (Fig. 1). Back mapping of long reads to these repeats verified the different conformations for each accession and that the varied conformations did indeed exist simultaneously. The existence of these various conformations and structural variations in *Morus* mitogenomes may potentially influence their evolutionary dynamics, and their ability to adapt to environmental changes, and can also offer increased flexibility for interactions with the nuclear genome.[70] Although the repeat-mediated recombination events combined with several gene transfer events from plastomes resulted in a dramatic genome structure rearrangement (Fig. 1; Supplemental Fig. S3), the remaining 30

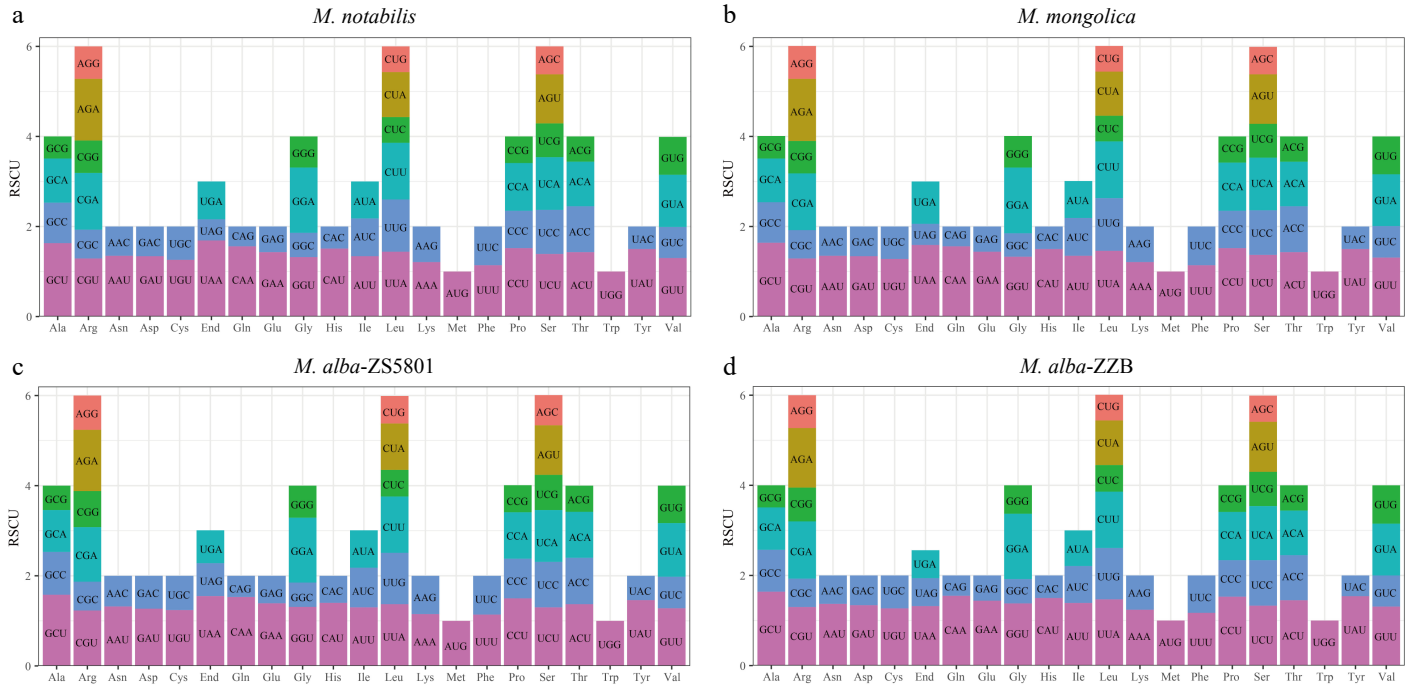

Fig. 4 Relative synonymous codon usage (RSCU) in the four *Morus* mitogenomes. (a) *M. notabilis*, (b) *M. mongolica*, (c) *M. alba*-ZS5801, and (d) *M. alba*-ZZB respectively.

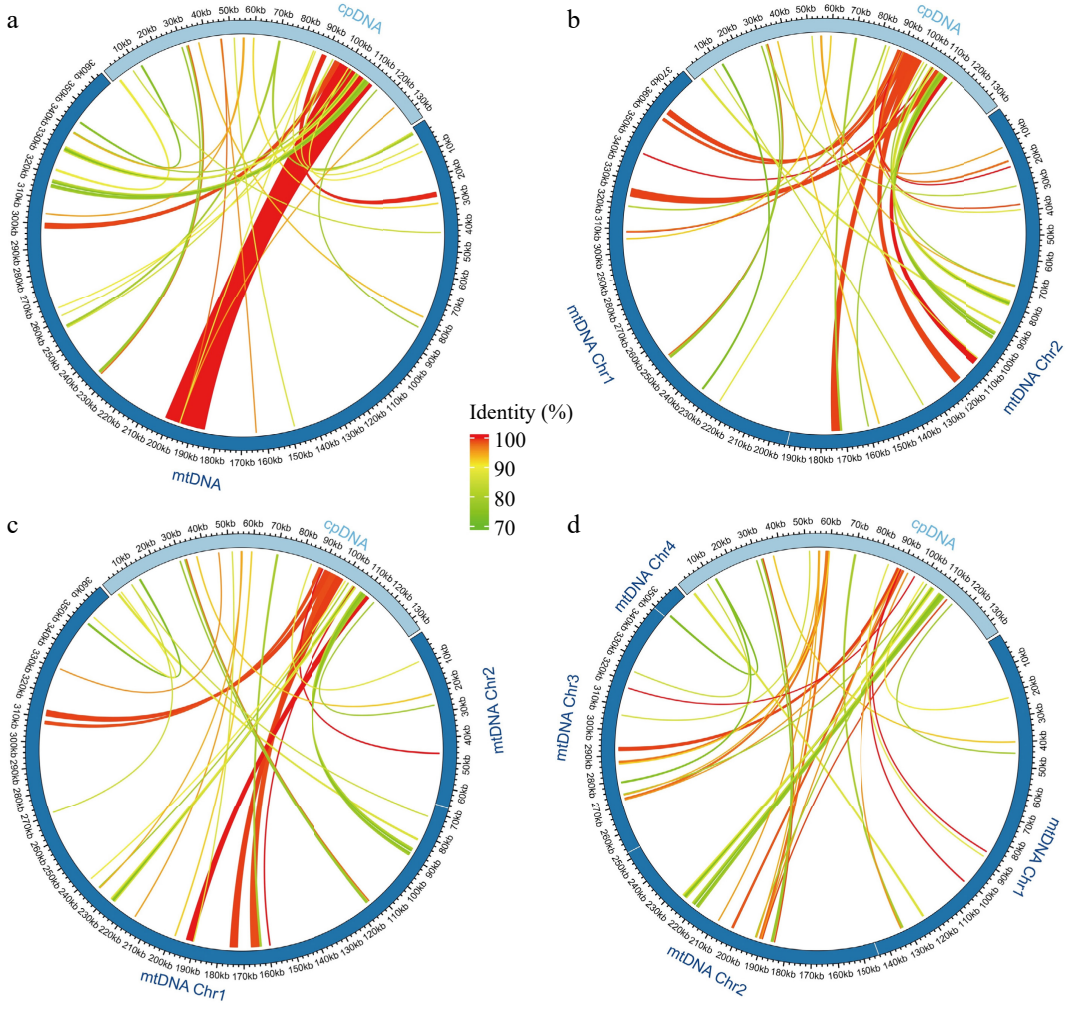

Fig. 5 Schematic representation of the distribution of MTPTs between plastome and mitogenomes for (a) *M. notabilis*, (b) *M. mongolica*, (c) *M. alba*-ZS5801, and (d) *M. alba*-ZZB. The MTPTs on the plastid IR regions were counted only once. Different colors of ribbons represent different identities. The length of each MTPT and the genes it contains can be found in Supplemental Table S6.

PCGs were shared across the four *Morus* mitogenomes with the exception of *rps*3 being lost in *M. alba*-ZZB and *rps*13 in *M. alba*-5801 (Supplemental Fig. S2). It is worth noting that the five exons of two PCGs including *nad*1 and *nad*5, which encode subunits 1 and 5 of the NADH-ubiquinone oxidoreductase complex were annotated into different genomic molecules, suggesting a layered recombination and gene loss mechanism for the creation of such unique mixed genome-derived chimeric genes. The mechanism by which these genes function normally would require deeper study.

Phylogenetic relationships of Rosales

For the past three decades, plant systematists have used DNA sequences to characterize and classify plant diversity, which have had a profound influence on our understanding of plant phylogenetic relationships and patterns of diversification[71]. With the increasing availability of genomic resources held in public databases and many newly developed pipelines for identifying orthologous genes (OrthoFinder, HybPiper)[72,73] or assembling plastomes (NOVOPlasty, GetOrganelle)[33,74], a large number of plastomes[75,76] and nuclear genes[5,21] have been used in phylogenomic inference for different plant lineages. Unlike plastomes or nuclear genes, the mitogenome is rarely used in phylogenetic inference mainly due to the slower substitutional rates of genes and difficult mitogenome assembly caused by frequent genomic rearrangements and foreign gene transfer[77,78]. However, the advent of third-generation sequencing and the development of fast and accurate software for assembling mitogenomes have promoted mitogenomes for phylogenetic analyses at higher taxonomic levels[24,64,65].

Here, the 40 species of Rosales including six of the nine families were selected for phylogenetic inference, representing the first use of mitogenomes to investigate the phylogenetic relationships of Rosales. In Rosales, the six families were divided into two clades based on ML and BI methods (Fig. 6). Two subfamilies of Rosaceae including Amygdaloideae and Rosoideae comprised one clade, which was consistent with the previous classification results of Rosaceae using plastid genes and nuclear genes[79,80]. However, the phylogenetic position of Dryadeae cannot be confirmed at the mitochondrial level due to the lack of samples from this subfamily. The remaining five families formed the other clade and the family Rhamnaceae was recovered as the first diverging lineage within it (Fig. 6), which was the most striking difference with previous studies. Whether phylogenetic analyses relied on a small number of plastid genes[81,82], whole plastomes[75,83], or even numerous nuclear

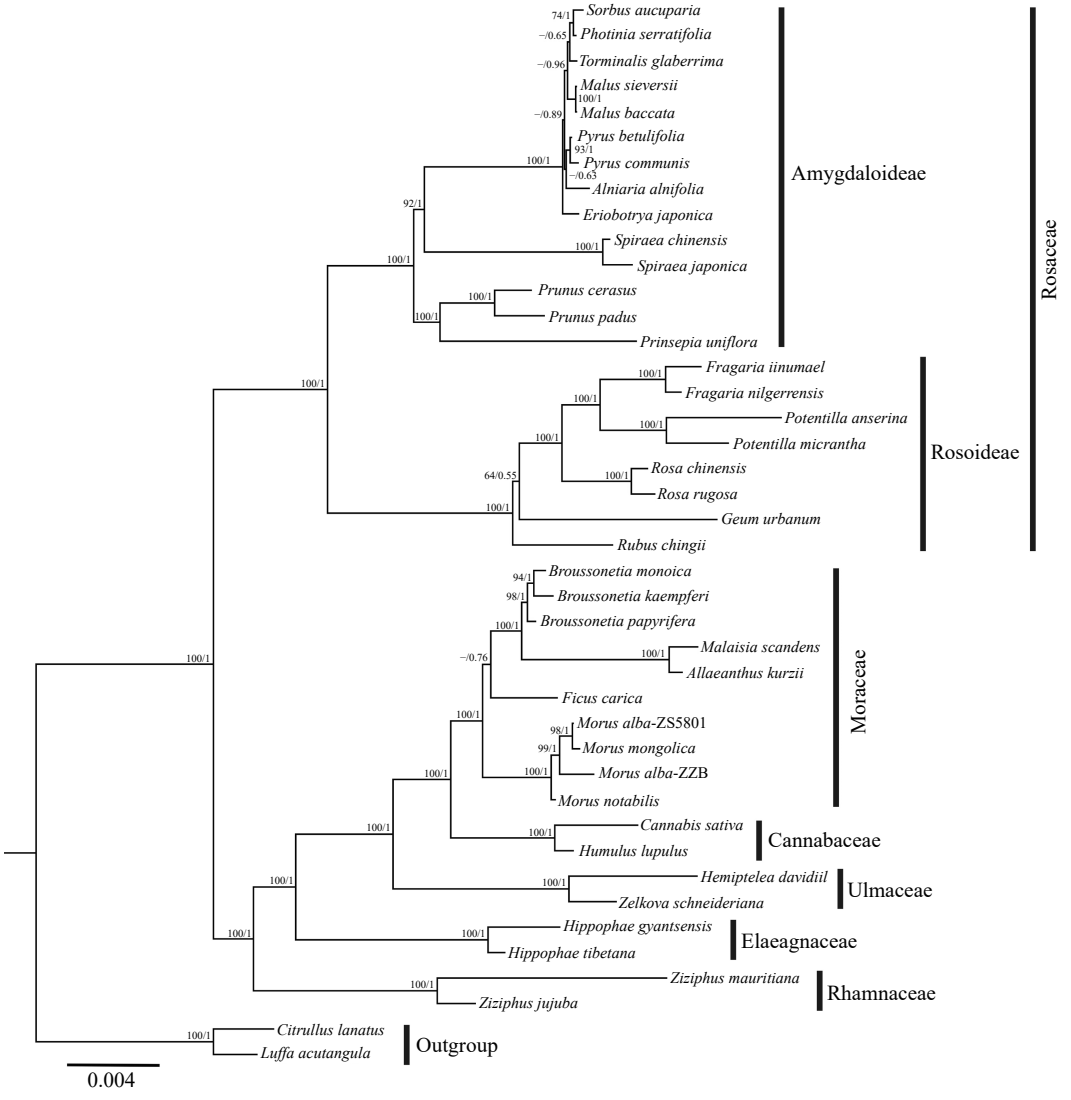

Fig. 6 Phylogenetic trees of Rosales estimated using RA × ML from 25 PCGs of the mitogenomes. Numbers above the branches represent ML bootstrap/Bayesian posterior probability (BS/PP), ML bootstrap values with lower than 50 are indicated by the dash.

genes[22,84], the phylogenetic analyses all showed that the family Rhamnaceae was resolved as a sister to the family Elaeagnaceae and the clade was subsequently sister to a clade including Moraceae, Cannabaceae and Ulmaceae. As is known, the plastid and mitogenomes typically are both maternally inherited in angiosperms[85,86], but the different placement of Rhamnaceae from mitochondrial and plastid data suggest that the evolutionary histories of these two subcellular compartments are unlinked. However, the biparental inheritance of organellar genomes has been documented[87,88], and paternal inheritance of plastomes has been reported in *Turnera ulmifolia*[89] and *Medicago sativa*[90], which were closely related to Rosales in the phylogenetic tree[21]. Therefore, it was deduced that the identical placement of Rhamnaceae in nuclear and plastid phylogenies due to the paternal inheritance of plastomes of this family, and the conflicting topologies of Rhamnaceae in plastid and mitochondrial phylogenies were mainly caused by ancient hybridization. The above speculation however needs to be further confirmed *via* expanded sampling of Rhamnaceae in future work. It is worth noting that the two cultivars of *M. alba*, 'Zhongsang5801' and 'Zhenzhubai', did not cluster together due to the embedding of *M. mongolica*, suggesting frequent hybridization in *Morus* species. This result was also recovered in the previous study based on plastome data[5].

Conclusions

In the present study, four high-quality mitogenomes of *M. notabilis*, *M. mongolica*, *M. alba* 'Zhongsang5801' and *M. alba* 'Zhenzhubai' were assembled based on a hybrid strategy, and various aspects explored including genome structure, gene transfer, and phylogenetic implications. The mitogenomes were structurally heterogeneous across four *Morus* accessions and multiple genome conformations existed simultaneously for each species. The gene content was relatively conserved especially for PCGs, but the repeat sequences and foreign gene transfer sites varied widely in the four mitogenomes. The phylogeny of Rosales was investigated based on mitogenome sequences, the phylogenetic position of Rhamnaceae showed a strong difference compared to earlier plastid and nuclear phylogenies, which provided us with new insights into the evolution of Rosales.

Author contributions

The authors confirm contribution to the paper as follows: conceptualization, project administration, resources, supervision: Liu L, Li P; data curation: Long Q, Qian J, Shi Y; formal analysis, investigation: Long Q, Lv W; methodology: Qian J, Shi Y; writing – original draft: Long Q, Lv W; writing - review and editing: Liu L, Egan AN, Li P; funding acquisition: Liu L. All authors reviewed the results and approved the final version of the manuscript.

Data availability

The raw sequencing reads of the HiFi and Illumina for *M. mongolica* mitogenome and plastome assembly have been deposited in the NCBI database with BioProject IDs PRJNA1141701 and PRJNA1140905, and the data are available under the accession numbers SRR30031178 and SRR30012719. The mitogenome and plastome sequences of the four *Morus* accessions have been submitted to the NCBI database, and the GenBank accession numbers for mitogenomes are shown in Supplemental Table S2, while the GenBank accession numbers for *M. notabilis*, *M. mongolica*, *M. alba*-ZS5801 and *M. alba*-ZZB are PQ083379, PQ083380, PQ083381, and PQ083382, respectively.

Acknowledgments

This work was financially supported by the Xinyang Academy of Ecological Research Open Foundation (Grant No. 2023DBS09), the Henan Province Science and Technology Research Project (NO. 232102320272), the Henan Province Major Research Fund of Public Welfare (Grant No. 201300110900), and the Key Technology Research and Development Program of Zhejiang Province (Grant No. 2023C03138).

Conflict of interest

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

Supplementary information accompanies this paper at (https://www.maxapress.com/article/doi/10.48130/gcomm-0024-0005)

Dates

Received 24 October 2024; Revised 15 November 2024; Accepted 20 November 2024; Published online 28 November 2024

References

1. Govaerts R, Nic Lughadha E, Black N, Turner R, Paton A. 2021. The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity. *Scientific Data* 8:215
2. Chang JC, Hsu YH. 2021. Mulberry. In *Temperate Fruits: Production, Processing, and Marketing*, eds. Mandal D, Wermund U, Phavaphutanon L, Cronje R. New York: Apple Academic Press. pp. 491–535. doi: 10.1201/9781003045861-9
3. Jain M, Bansal J, Rajkumar MS, Sharma N, Khurana JP, et al. 2022. Draft genome sequence of Indian mulberry (*Morus indica*) provides a resource for functional and translational genomics. *Genomics* 114:110346
4. Dai F, Zhuo X, Luo G, Wang Z, Xu Y, et al. 2023. Genomic resequencing unravels the genetic basis of domestication, expansion, and trait improvement in *Morus atropurpurea*. *Advanced Science* 10:e2300039
5. Wang M, Zhu M, Qian J, Yang Z, Shang F, et al. 2024. Phylogenomics of mulberries (*Morus*, Moraceae) inferred from plastomes and single copy nuclear genes. *Molecular Phylogenetics and Evolution* 197:108093
6. Jiao F, Luo R, Dai X, Liu H, Yu G, et al. 2020. Chromosome-level reference genome and population genomic analysis provide insights into the evolution and improvement of domesticated mulberry (*Morus alba*). *Molecular Plant* 13:1001–12
7. Xia Z, Dai X, Fan W, Liu C, Zhang M, et al. 2022. Chromosome-level genomes reveal the genetic basis of descending Dysploidy and sex determination in *Morus* plants. *Genomics, Proteomics & Bioinformatics* 20:1119–37
8. Kim BS, Kim H, Kang SS. 2019. *In vitro* anti-bacterial and anti-inflammatory activities of lactic acid bacteria-biotransformed mulberry (*Morus alba* Linnaeus) fruit extract against *Salmonella* Typhimurium. *Food Control* 106:106758
9. Suriyaprom S, Srisai P, Intachaisri V, Kaewkod T, Pekkoh J, et al. 2023. Antioxidant and anti-inflammatory activity on LPS-stimulated RAW 264.7 macrophage cells of white mulberry (*Morus alba* L.) leaf extracts. *Molecules* 28:4395
10. Donno D, Mellano MG, Gamba G, Riondato I, Beccaro GL. 2021. Mulberry: an ornamental tree that gives bioactive compounds for human health. *Acta Horticulturae* 1331:205–14
11. Linnaeus C. 1753. *Morus*. *Species Plantarum* 2:968
12. Zhao W, Pan Y, Zhang ZF, Jia SH, Miao X, et al. 2005. Phylogeny of the genus *Morus* (Urticales: Moraceae) inferred from ITS and *trn*L-F sequences. *African Journal of Biotechnology* 4:563–69

13. Burgess KS, Morgan M, Deverno L, Husband BC. 2005. Asymmetrical introgression between two *Morus* species (*M*. alba, M. rubra) that differ in abundance. *Molecular Ecology* 14:3471–83
14. Burgess KS, Morgan M, Husband BC. 2008. Interspecific seed discounting and the fertility cost of hybridization in an endangered species. *New Phytologist* 177:276–84
15. Das B, Krishnaswami S. 1965. Some observations on interspecific hybridization in mulberry. *Inidan Journal of Sericulture* 4:1–8
16. Nepal MP, Purintun JM. 2021. Systematics of the genus *Morus* L. (Moraceae): taxonomy, phylogeny and potential responses to climate change. In *Mulberry: Genetic Improvement in Context of Climate Change*, eds. Razdan MK, Thomas D. Boca Raton: CRC Press. pp. 2–20. doi: 10.1201/9780429399237-2
17. Tani N, Kawahara T, Yoshimaru H, Hoshi Y. 2003. Development of SCAR markers distinguishing pure seedlings of the endangered species *Morus boninensis* from *M. boninensis* × *M. acidosa* hybrids for conservation in Bonin (Ogasawara) Islands. *Conservation Genetics* 4:605–12
18. Nepal MP, Ferguson CJ. 2012. Phylogenetics of *Morus* (Moraceae) inferred from ITS and *trnL-trnF* sequence data. *Systematic Botany* 37:442–50
19. Li HT, Yi TS, Gao LM, Ma PF, Zhang T, et al. 2019. Origin of angiosperms and the puzzle of the Jurassic gap. *Nature Plants* 5:461–70
20. Daniell H, Jin S, Zhu XG, Gitzendanner MA, Soltis DE, et al. 2021. Green giant—a tiny chloroplast genome with mighty power to produce high-value proteins: history and phylogeny. *Plant Biotechnology Journal* 19:430–47
21. Liu L, Chen M, Folk RA, Wang M, Zhao T, et al. 2023. Phylogenomic and syntenic data demonstrate complex evolutionary processes in early radiation of the rosids. *Molecular Ecology Resources* 23:1673–88
22. Zuntini AR, Carruthers T, Maurin O, Bailey PC, Leempoel K, et al. 2024. Phylogenomics and the rise of the angiosperms. *Nature* 629:843–50
23. Birky CW Jr. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences* 92:11331–38
24. Zhang X, Li P, Wang J, Fu D, Zhao B, et al. 2024. Comparative genomic and phylogenetic analyses of mitochondrial genomes of hawthorn (*Crataegus* spp.) in Northeast China. *International Journal of Biological Macromolecules* 272:132795
25. Mower JP, Sloan DB, Alverson AJ. 2012. Plant mitochondrial genome diversity: the genomics revolution. *Plant genome diversity: plant genomes, their residents, and their evolutionary dynamics*, eds. Wendel J, Greilhuber J, Dolezel J, Leitch I. Volume 1. Vienna: Springer. pp. 123–44. doi: 10.1007/978-3-7091-1130-7_9
26. Skippington E, Barkman TJ, Rice DW, Palmer JD. 2015. Miniaturized mitogenome of the parasitic plant *Viscum scurruloideum* is extremely divergent and dynamic and has lost all *nad* genes. *Proceedings of the National Academy of Sciences of the United States of America* 112:E3515–E3524
27. Putintseva YA, Bondar EI, Simonov EP, Sharov VV, Oreshkova NV, et al. 2020. Siberian larch (*Larix sibirica* Ledeb.) mitochondrial genome assembled using both short and long nucleotide sequence reads is currently the largest known mitogenome. *BMC Genomics* 21:654
28. Cheng N, Lo YS, Ansari MI, Ho KC, Jeng ST, et al. 2017. Correlation between mtDNA complexity and mtDNA replication mode in developing cotyledon mitochondria during mung bean seed germination. *New Phytologist* 213:751–63
29. Kozik A, Rowan BA, Lavelle D, Berke L, Schranz ME, et al. 2019. The alternative reality of plant mitochondrial DNA: One ring does not rule them all. *PLoS Genetics* 15:e1008373
30. Maréchal A, Brisson N. 2010. Recombination and the maintenance of plant organelle genome stability. *New Phytologist* 186:299–317
31. Jiang L. 2020. Male sterility in maize: A precise dialogue between the mitochondria and nucleus. *Molecular Plant* 13:1237
32. Garcia LE, Edera AA, Palmer JD, Sato H, Sanchez-Puerta MV. 2021. Horizontal gene transfers dominate the functional mitochondrial gene space of a holoparasitic plant. *New Phytologist* 229:1701–14
33. Jin JJ, Yu WB, Yang JB, Song Y, DePamphilis CW, et al. 2020. GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biology* 21:241
34. He W, Xiang K, Chen C, Wang J, Wu Z. 2023. Master graph: an essential integrated assembly model for the plant mitogenome based on a graph-based framework. *Briefings in Bioinformatics* 24(1):bbac522
35. Bi C, Shen F, Han F, Qu Y, Hou J, et al. 2024. PMAT: an efficient plant mitogenome assembly toolkit using low-coverage HiFi sequencing data. *Horticulture Research* 11:uhae023
36. Guo W, Grewe F, Fan W, Young GJ, Knoop V, et al. 2016. *Ginkgo* and *Welwitschia* mitogenomes reveal extreme contrasts in gymnosperm mitochondrial evolution. *Molecular Biology and Evolution* 33:1448–60
37. Vargas OM, Ortiz EM, Simpson BB. 2017. Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: Astereae: *Diplostephium*). *New Phytologist* 214:1736–50
38. Liu BB, Ma ZY, Ren C, Hodel RGJ, Sun M, et al. 2021. Capturing single-copy nuclear genes, organellar genomes, and nuclear ribosomal DNA from deep genome skimming data for plant phylogenetics: a case study in Vitaceae. *Journal of Systematics and Evolution* 59:1124–38
39. Liu LX, Deng P, Chen MZ, Yu LM, Lee J, et al. 2023. Systematics of *Mukdenia* and *Oresitrophe* (Saxifragaceae): Insights from genome skimming data. *Journal of Systematics and Evolution* 61:99–114
40. Wang J, Kan S, Liao X, Zhou J, Tembrock LR, et al. 2024. Plant organellar genomes: much done, much more to do. *Trends in Plant Science* 29:754–69
41. Chen Y, Ye W, Zhang Y, Xu Y. 2015. High speed BLASTN: an accelerated MegaBLAST search tool. *Nucleic Acids Research* 43:7762–68
42. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31:3350–52
43. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–49
44. Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Research* 33:W686–W689
45. Lohse M, Drechsel O, Kahlau S, Bock R. 2013. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Research* 41:W575–W581
46. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–60
47. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–79
48. Milne I, Bayer M, Cardle L, Shaw P, Stephen G, et al. 2010. Tablet—next generation sequence assembly visualization. *Bioinformatics* 26:401–2
49. Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* 14:178–92
50. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, et al. 2020. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Molecular Ecology Resources* 20:348–55
51. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J et al. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research* 29:4633–42
52. Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* 27:573–80
53. Beier S, Thiel T, Münch T, Scholz U, Mascher M. 2017. MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33:2583–85
54. Zhang H, Meltzer P, Davis S. 2013. RCircos: an R package for Circos 2D track plots. *BMC Bioinformatics* 14:244
55. Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* 34:772–73
56. Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology* 57:758–71
57. Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–74

58. Davila JI, Arrieta-Montiel MP, Wamboldt Y, Cao J, Hagmann J, et al. 2011. Double-strand break repair processes drive evolution of the mitochondrial genome in Arabidopsis. *BMC Biology* 9:64
59. Cole LW, Guo W, Mower JP, Palmer JD. 2018. High and variable rates of repeat-mediated mitochondrial genome rearrangement in a genus of plants. *Molecular Biology and Evolution* 35:2773–85
60. Odahara M, Nakamura K, Sekine Y, Oshima T. 2021. Ultra-deep sequencing reveals dramatic alteration of organellar genomes in *Physcomitrella patens* due to biased asymmetric recombination. *Communications Biology* 4:633
61. Wang S, Li D, Yao X, Song Q, Wang Z, et al. 2019. Evolution and diversification of kiwifruit mitogenomes through extensive whole-genome rearrangement and mosaic loss of intergenic sequences in a highly variable region. *Genome Biology and Evolution* 11:1192–206
62. Ogihara Y, Yamazaki Y, Murai K, Kanno A, Terachi T, et al. 2005. Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. *Nucleic Acids Research* 33:6235–50
63. Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, et al. 2012. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biology* 10:e1001241
64. Li J, Li J, Ma Y, Kou L, Wei J, et al. 2022. The complete mitochondrial genome of okra (*Abelmoschus esculentus*): using nanopore long reads to investigate gene transfer from chloroplast genomes and rearrangements of mitochondrial DNA molecules. *BMC Genomics* 23:481
65. Lai C, Wang J, Kan S, Zhang S, Li P, et al. 2022. Comparative analysis of mitochondrial genomes of *Broussonetia* spp. (Moraceae) reveals heterogeneity in structure, synteny, intercellular gene transfer, and RNA editing. *Frontiers in Plant Science* 13:1052151
66. Wei L, Zheng T, Xiang J, Cheng J, Wu J. 2023. Assembly and analysis of the first complete mitochondrial genome of *Ficus carica* Linn. *Journal of the American Society for Horticultural Science* 148:283–91
67. Palmer JD, Adams KL, Cho Y, Parkinson CL, Qiu YL, et al. 2000. Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. *Proceedings of the National Academy of Sciences of the United States of America* 97:6960–66
68. Park S, Grewe F, Zhu A, Ruhlman TA, Sabir J, et al. 2015. Dynamic evolution of *Geranium* mitochondrial genomes through multiple horizontal and intracellular gene transfers. *New Phytologist* 208:570–83
69. Yang Z, Ni Y, Lin Z, Yang L, Chen G, et al. 2022. De novo assembly of the complete mitochondrial genome of sweet potato (*Ipomoea batatas* [L.] Lam) revealed the existence of homologous conformations generated by the repeat-mediated recombination. *BMC Plant Biology* 22:285
70. Yang M, Zhang S, Li B, Yang Y, Lan Y, et al. 2024. Comprehensive analysis of the mitochondrial genome of *Ziziphus mauritiana* Lam.: Investigating sequence structure, repeat-mediated recombination, chloroplast-derived transfer, RNA editing, and evolution. *Scientia Horticulturae* 324:112637
71. Straub SCK, Parks M, Weitemier K, Fishbein M, Cronn RC, et al. 2012. Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *American Journal of Botany* 99:349–64
72. Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology* 16:157
73. Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, et al. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* 4:1600016
74. Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Research* 45:e18
75. Li HT, Luo Y, Gan L, Ma PF, Gao LM, et al. 2021. Plastid phylogenomic insights into relationships of all flowering plant families. *BMC Biology* 19:232
76. Hu H, Sun P, Yang Y, Ma J, Liu J. 2023. Genome-scale angiosperm phylogenies based on nuclear, plastome, and mitochondrial datasets. *Journal of Integrative Plant Biology* 65:1479–89
77. Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America* 84:9054–58
78. Rice DW, Alverson AJ, Richardson AO, Young GJ, Sanchez-Puerta MV, et al. 2013. Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. *Science* 342:1468–73
79. Potter D, Eriksson T, Evans RC, Oh S, Smedmark JEE, et al. 2007. Phylogeny and classification of Rosaceae. *Plant Systematics and Evolution* 266:5–43
80. Xiang Y, Huang CH, Hu Y, Wen J, Li S, et al. 2017. Evolution of rosaceae fruit types based on nuclear phylogeny in the context of geological times and genome duplication. *Molecular Biology & Evolution* 34:262–81
81. Chen ZD, Yang T, Lin L, Lu LM, Li HL, et al. 2016. Tree of life for the genera of Chinese vascular plants. *Journal of Systematics and Evolution* 54:277–306
82. Sun M, Naeem R, Su JX, Cao ZY, Burleigh JG, et al. 2016. Phylogeny of the *Rosidae*: A dense taxon sampling analysis. *Journal of Systematics and Evolution* 54:363–91
83. Shi C, Han K, Li L, Seim I, Lee SMY, et al. 2020. Complete chloroplast genomes of 14 mangroves: phylogenetic and comparative genomic analyses. *BioMed Research International* 2020:8731857
84. xxx. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574:679–685
85. Mogensen HL. 1996. INVITED SPECIAL PAPER: the hows and whys of cytoplasmic inheritance in seed plants. *American Journal of Botany* 83:383–404
86. Birky CW Jr . 2001. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annual Review of Genetics* 35:125–48
87. Havey MJ, McCreight JD, Rhodes B, Taurick G. 1998. Differential transmission of the *Cucumis* organellar genomes. *Theoretical and Applied Genetics* 97:122–28
88. Yang TW, Yang YA, Xiong Z. 2000. Paternal inheritance of chloroplast DNA in interspecific hybrids in the genus *Larrea* (Zygophyllaceae). *American Journal of Botany* 87:1452–58
89. Shore JS, McQueen KL, Little SH. 1994. Inheritance of plastid DNA in the *Turnera ulmifolia* complex (Turneraceae). *American Journal of Botany* 81:1636–39
90. Masoud SA, Johnson LB, Sorensen EL. 1990. High transmission of paternal plastid DNA in alfalfa plants demonstrated by restriction fragment polymorphic analysis. *Theoretical and Applied Genetics* 79:49–55