The Complete Mitochondrial Genome Sequence Variation and Phylogenetic Analysis of Mulberry

Guo Liangliang
Jiangsu University of Science and Technology

Shi Yisu
Jiangsu University of Science and Technology

Wu Mengmeng
Jiangsu University of Science and Technology

Michael Ackah
Jiangsu University of Science and Technology

Guo Peng
Jiangsu University of Science and Technology

Zheng Danyan
Jiangsu University of Science and Technology

Qiang Lin
Guangxi Zhuang Autonomous Regin, Nanning, 53007

Weiguo Zhao ( wgzsri@126.com )
Jiangsu University of Science and Technology

Keywords: M. multicaulis, M. atropurpurea, Mitochondrial genome, variation, Phylogenetic analysis

DOI: https://doi.org/10.21203/rs.3.rs-445726/v1

License: ☑️ ① This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Mulberry is admired for its landscaping and possesses high development prospects and scientific research value. Mitochondria are the plants' powerhouse that produces energy to carry out life processes. In this study, the mt genome of *Morus* L (*M. atropurpurea* and *M. multicaulis*) were annotated and assembled. The circular mt genome of *M. multicaulis* has a length of 361,546bp, contains 54 genes, including 31 protein-coding genes, 20 tRNA genes, and 3 rRNA genes and composition of A (27.38%), T (27.20%), C (22.63%) and G (22.79%). The sequence repeats, RNA editing gene and migration from cp to mt and was observed in *M. multicaulis* mt genome. Phylogenetic analysis based on the complete mt genomes of *Morus* and other 28 species reects an exact evolutionary and taxonomic status.

Furthermore, we investigation on mt genome size, organization, and plastomes at the global level and pi analysis of Morus genome was investigated and compared to other land plants. The results indicate that the exist mt genome's variation in plants. We reported the mt genome assembly and annotation of a halophytic model plant, *M. multicaulis*, and subsequent analysis, which provided us with a comprehensive understanding of the *Morus* mt genome.

Introduction

*Morus* is an economically significant crop and strong saline-alkali tolerance belonging to the Moraceae family which is native to China and also been planted in various areas for erosion control and windbreaks all over the world (He N, 2013). Great value as a food for healthy *M. alba* also can be used as a herbal medicine to cure fever, improve eyesight, strengthen joints, and lower blood pressure in China (Chan et al., 2016).

The mitochondrial (mt) genome is power source for energy synthesis and conversion, providing energy protection for various physiological activities of cells (Kozik, Rowan, Lavelle, Berke, & Christensen, 2019) such as cell differentiation, apoptosis, cell growth and cell division (Rehman et al., 2012). In addition, it is also involved in the synthesis and degradation of several compounds (Shtolz, Dan, & Evolution, 2019), therefore, mitochondrial play an essential role in plant productivity and development (Yasunari et al., 2005). The mt genome with highly conserved, but the mt genomes has significant differences in length, gene sequence and content (Richardson, Rice, Young, Alverson, & Palmer, 2013). The smallest known terrestrial plant is about 66 Kb, and the largest terrestrial plant mt genome length is 11.3 Mb (Daniel et al., 2012; Skippington, Barkman, Rice, & Palmer, 2015), mostly vary from 200 kb to 3Mb and larger than mt genomes of other eukaryotes (X. & Physiology, 2006). The mt genome structures are shaped by active recombination, gene transfer to the nucleus, and other forces that remain unclear show that by Physical mapping and sequencing of some of the small mt genomes (Woloszynska, 2009). Structural analyses revealed high frequencies of intra- and intermolecular recombination, which generated a structurally dynamic assemblage of genome configurations (Alverson et al., 2010). The mt genome are inherited from the maternal parent (Wolfe, Li, & Sharp, 1988), this provides a powerful model for the study of genome structure and evolution, also a certain advantages in phylogenetic reconstruction. These genomes exhibit an intriguing mixture of conservative (slowest rates of nucleotide substitution) (Drouin, Daoud, Xia, & Evolution, 2008) and dynamic evolutionary patterns. Some previous reported (Tong, Kim, & Park, 2016) also suggested that for evolutionary studies it is not necessary to assemble whole organelle genomes but just exploring the variations.

Currently, With the rapid development of sequencing technology, an increasing number of complete plant mt genomes were assembled and reported Up to Jan. 2021, 351 complete mt genomes have been deposited in GenBank Organelle Genome Resources (Cheng, He, Priyadarshani, Wang, & Qin, 2021). However, the mt genome of *Morus* is incomplete and unexplored. In this study, we sequenced and annotated the mt genome of cultivated *Morus* (*M. atropurpurea* and *M. multicaulis*) and compared it with the wild *M. notabilis* (NC-041177.1) and other eudicot which provides additional information for a better understanding of the genetics of the *Morus* L.

Material And Methods

Plant material, DNA extraction, and sequencing

The *M. atropurpurea* and *M. multicaulis* plants were collected from Jiangsu University of Science and Technology Sericulture Research Institute. National Mulberry Genebank Zhenjiang, China affiliated to the Chinese Academy Of Agricultural Sciences. Materials used this time have been officially approved by Chinese Academy Of Agricultural Sciences. Plant Genomic DNA Kit was used to isolate total genomic DNA from 100 mg fresh leaves and DNA sample quality was examined with agarose-gel electrophoresis, and the concentration was measured by Nanodrop instrument then qualified samples were sent to the Oxford Nanopore PromethION for sequencing.

Assembly and annotation of the mitochondrial genome
The mt genome sequence of mulberry were selected using blast v2.6 (https://blast.ncbi.nlm.nih.gov/Blast.cgi) align the contig with the plant mitochondrial gene database (the mitochondrial gene sequence of the species published on NCBI). Subsequently, assembled by the software Canu (Sergey et al., 2017) with the selected reads. Use NextPolish1.3.1 (https://github.com/Nextomics/NextPolish) to calibrate and pilon (Walker, 2015) correct read errors to get the final assembly results. The encoded protein and rRNA use blast to align the published plant mitochondrial sequence as a ref, and then make further manual adjustments according to relative species. TRNA is annotated with tRNAscanSE (http://lowelab.ucsc.edu/tRNAscan-SE/). ORF uses OpenReading Frame Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) for annotation. After checking and manually confirmed the final annotation result is obtained. Use OGDRAW (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html) to draw a circular mitochondrial genome map.

Analysis of repeated sequences

The software misa (1.0) was identification tool used to detect simple sequence repeats (Sebastian et al., 2017). The type of p1, p2, p3, p4, p5, and p6 refer to (1/8) (2/5) (3/3) (4/3) (5/3) (6/3) bases with (unit size / minimum number of repeats), were identified in this analysis. The scattered repetitive sequences were detected using vmatch v2.3.0 (http://www.vmatch.de/) Combining Perl scripts to identify repetitive sequences, with minimum length set to 30 bp and hamming distance was 3.

RNA editing analyses and chloroplast to mitochondrion DNA transformation

The editing sites in the mitochondrial RNA of *M. multicaulis* were use online sites to predictions (http://www.prepact.de/prepact-main.php). DNA migration is common in plants and varies from species to species occurs during autophagy, gametogenesis, and fertilization (Huang, Ayliffe, & Timmis, 2003). We using blast software set similarity to 70%, and E-value to 10E-5 use circos v0.69-5 to visualize it. The cpDNA of *M. multicaulis* (KU355297) was downloaded from NCBI Organelle Genome Resources Database.

Variation architecture and Phylogenetic tree construction

Nucleic acid diversity (pi) by maft software (--auto mode) to compare the homologous gene sequences of different species globally, and use dnasp5 to calculate. Comparison of the mt genome sequence with other plastomes at the global level using mVISTA online software in shuffle-LAGAN mode. MEGA7.0 was used for phylogenetic tree construction by the maximum likelihood (ML) and neighbor-joining (NJ) methods with a bootstrap of 1,000 Poisson model. The mt genomes date were downloaded from NCBI, including *M. notabilis* (NC-041177.1).

Statement

The experimental materials for this time are only experimental research and no field research. They were collected from my own school, National Mulberry GeneBank Zhenjiang. This collection was supported by Chinese Academy Of Agricultural Sciences and with the guidance of the school official leaders. The collection was conducted under the conditions permitted by national laws and regulations, strictly abide by relevant laws and get official permission. My corresponding author also made a statement. After the collection, it is only for experimental research and has no other purpose.

Results

Genome content and organization

The *M. multicaulis* mt genome is circular was determined to be 361,546 bp long, base composition of the genome is A (27.38%), T (27.20%), C (22.63%), G (22.79%) contains 54 functional genes including 3 rRNA genes, 20 tRNA genes, and 31 PCGs Pseudogenes and ORFs were all non-coding (Table 1). The mt genome of *M. multicaulis* functional categorization and physical locations of the annotated genes were presented (Fig. 1), encodes 31 different protein that could be divided into 9 classes (Table 3): ATP Synthase (5 genes), Cytochrome C Biogenesis (4 genes), Ubiquinol Cytochrome c Reductase (1 gene), Cytochrome C oxidase (3 genes), Maturases (2 gene), Transport membrane protein (1 gene), NADH dehydrogenase (9 genes), Ribosomal proteins (SSU) (5 genes) and Ribosomal proteins (LSU) (1 gene).

The mt genome sequence of *M. atropurpurea* is also circular was to be found 395,412-bp long (Figure 2) base composed of C+G (45.50%) including 57 functional genes contains 2 rRNA genes, 22 tRNA genes and 32 PCGs, 31 different protein slao can be divided into 9 classes (Table 2).
### Table 1

| Characteristics | *M. notabilis* | *M. multicaulis* | *M. atropurpurea* |
|-----------------|---------------|------------------|-------------------|
| Size (bp)       | 362,069       | 361,546          | 395,412           |
| GC content (%)  | 45.66         | 45.42            | 45.50             |
| Number of genes | 54            | 54               | 57                |
| Protein-coding genes | 26        | 31               | 32                |
| rRNA            | 3             | 3                | 3                 |
| tRNA            | 21            | 20               | 22                |

### Variations and Codon usage

In this study, the mt genome of *M. multicaulis* and *M. atropurpurea* were encoded by 27,933 and 28,251 codons. For *M. multicaulis* 31 protein-coding genes in the mt genome were encoded by 27,933 codons respectively, 62.2% of codons end in A or T. Leu accounts for the highest codon usage (3,084) followed by Ser (2,454) and Arg (1,824) (Fig. 3). These three amino acids almost represent four fifths of the total codons. The least number of codons is Trp (459). All of the protein-coding genes used AUG (753) most common start codon and three stop codons UAA, UGA, and UAG with the following utilization rate: UAA (53.33%), UGA (23.33%), and UAG (23.33%), (Table 3). Interesting for *M. atropurpurea* the most high codon usage is Leu and follow by Ser and Arg (Fig. 4).

Previous results shown that the mt genomes contain variable number of introns (Xiaofang et al., 2018). In our results the mt genome of *M. multicaulis* has 8 intron-containing genes (ccmFC, cox2, nad1, nad2, nad4, nad5, nad7, trnF-AAA) harboring 21 introns in total and nad1, nad2, nad5, nad7 even contains 4 introns, which is the highest intron number. *M. atropurpurea* also have 8 intron-containing genes contain 21 introns. Most land plants contain 3 rRNA genes (Archibald, 2011). Consistently, our two species three same rRNA genes rrn18, rrn26 and rrn5 were annotated in *Morus* mt genome. Besides, 20 different transfer RNAs were identified in *M. multicaulis* mt genome transporting 19 amino acids, which more than one transfer RNAs might transport the same amino acid with different codons (Table 3).
| Group of genes                  | *M. multicaulis*                  | *M. atropurpurea*                  |
|--------------------------------|----------------------------------|-----------------------------------|
| ATP synthase                   | atp1 atp4 atp6 atp8 atp9         | atp1 atp4 atp6 atp8 atp9          |
| Cytochrome c biogenesis        | ccmB ccmC ccmFC* ccmFN           | ccmB ccmC ccmFC* ccmFN            |
| Ubichinol cytochrome creductase| cob                              | Cob                               |
| Cytochrome c oxidase           | cox1 cox2* cox3                  | cox1 cox2* cox3                   |
| Maturases                      | matR(2)                          | matR                              |
| Transport membrane protein     | mttB                             | mttB                              |
| NADH dehydrogenase             | nad1**** nad2**** nad3 nad4** nad4L nad5**** nad6 nad7**** nad9 | nad1**** nad2**** nad3 nad4** nad4L nad5**** nad6 nad7**** nad9 |
| Ribosomal proteins (LSU)       | rpl16                            | rpl16                             |
| Ribosomal proteins (SSU)       | rps12 rps19 rps3 rps4 rps7       | rps12 rps13 rps19(2) rps3 rps4 rps7 |
| Succinate dehydrogenase        | ψsdh4                            | ψsdh4                             |
| Ribosomal RNAs                 | rm18 rm26 rm5                    | rm18 rm26 rm5                     |
| Transfer RNAs                  | trmC-GCA trmD-GTC trmE-TTC trmF-AAA* trmF-GAA trmK-TTT trmL-CAA trmM-CAT(4) trmN-GTT trmP-TGG(2) trmQ-TTG(2) trmR-ACG trmS-TGA trmW-CCA trmY-GTA | trmA-TGC*(2) trmC-GCA trmD-GTC trmE-TTC(2) trmF-AAA* trmF-GAA trmK-TTT trmL-CAA trmM-CAT(4) trmN-GTT trmP-TGG(2) trmQ-TTG trmR-ACG trmS-TGA trmW-CCA trmY-GTA |

**Notes:** The numbers after the gene names indicate the duplication number. "**" indicate genes containing one or more introns. "ψ" indicate pseudogene.
| Codon | AA | M. multicaulis No. | RSCU | M. atropurpurea No. | RSCU | Codon | AA | M. multicaulis No. | RSCU | M. atropurpurea No. | RSCU |
|-------|----|-------------------|------|---------------------|------|-------|----|-------------------|------|---------------------|------|
| UAA   | Ter| 16                | 1.5999 | 15                  | 1.4517 | AUG   | Met| 251              | 1     | 256                | 1.9922 |
| UAG   | Ter| 7                 | 0.6999 | 6                   | 0.5805 | AAC   | Asn| 101              | 0.6756 | 102                | 0.6868 |
| UGA   | Ter| 7                 | 0.6999 | 10                  | 0.9678 | AAU   | Asn| 198              | 1.3244 | 195                | 1.3132 |
| GCA   | Ala| 148               | 0.9656 | 150                 | 0.9676 | CCA   | Pro| 135              | 1.0908 | 134                | 1.0828 |
| GCC   | Ala| 135               | 0.8808 | 139                 | 0.8968 | CCC   | Pro| 103              | 0.8324 | 104                | 0.8404 |
| GCG   | Ala| 75                | 0.4892 | 75                  | 0.484  | CCG   | Pro| 69               | 0.5576 | 69                 | 0.5576 |
| GCU   | Ala| 255               | 1.664  | 256                 | 1.6516 | CCU   | Pro| 188              | 1.5192 | 188                | 1.5192 |
| UGC   | Cys| 52                | 0.7592 | 54                  | 0.7714 | CAA   | Gln| 198              | 1.5408 | 206                | 1.5488 |
| UGU   | Cys| 85                | 1.2408 | 86                  | 1.2286 | CAG   | Gln| 59               | 0.4592 | 60                 | 0.4512 |
| GAC   | Asp| 100               | 0.6802 | 103                 | 0.6822 | AGA   | Arg| 138              | 1.362  | 138                | 1.3488 |
| GAU   | Asp| 194               | 1.3198 | 199                 | 1.3178 | AGG   | Arg| 75               | 0.7404 | 78                 | 0.762  |
| GAA   | Glu| 277               | 1.406  | 286                 | 1.4158 | CGA   | Arg| 124              | 1.2234 | 125                | 1.2216 |
| GAG   | Glu| 117               | 0.594  | 118                 | 0.5842 | CGC   | Arg| 66               | 0.6516 | 67                 | 0.6546 |
| UUC   | Phe| 263               | 0.858  | 262                 | 0.8576 | CGG   | Arg| 79               | 0.7794 | 77                 | 0.7524 |
| UUU   | Phe| 350               | 1.142  | 349                 | 1.1424 | CGU   | Arg| 126              | 1.2432 | 129                | 1.2606 |
| GGA   | Gly| 231               | 1.4236 | 238                 | 1.4448 | AGC   | Ser| 81               | 0.594  | 82                 | 0.5952 |
| GGC   | Gly| 88                | 0.5424 | 87                  | 0.528  | AGU   | Ser| 145              | 1.0638 | 148                | 1.074  |
| GGG   | Gly| 115               | 0.7088 | 116                 | 0.704  | UCA   | Ser| 157              | 1.1514 | 161                | 1.1682 |
| GGU   | Gly| 215               | 1.3252 | 218                 | 1.3232 | UCC   | Ser| 137              | 1.005  | 137                | 0.9942 |
| CAC   | His| 63                | 0.5338 | 64                  | 0.529  | UCG   | Ser| 106              | 0.7776 | 106                | 0.7692 |
| CAU   | His| 173               | 1.4662 | 178                 | 1.471  | UCU   | Ser| 192              | 1.4082 | 193                | 1.4004 |
| AUA   | Ile| 200               | 0.822  | 203                 | 0.8208 | ACA   | Thr| 118              | 0.9896 | 118                | 0.9792 |
| AUC   | Ile| 206               | 0.8466 | 207                 | 0.837  | ACC   | Thr| 123              | 1.0316 | 123                | 1.0208 |
| AUU   | Ile| 324               | 1.3314 | 332                 | 1.3422 | ACG   | Thr| 68               | 0.5704 | 68                 | 0.5644 |
| AAA   | Lys| 227               | 1.201  | 236                 | 1.2164 | ACU   | Thr| 168              | 1.4088 | 173                | 1.4356 |
| AAG   | Lys| 151               | 0.799  | 152                 | 0.7836 | GUA   | Val| 171              | 1.1592 | 171                | 1.1516 |
| CUA   | Leu| 169               | 0.9864 | 168                 | 0.9768 | GUC   | Val| 108              | 0.7324 | 109                | 0.734  |
| CUC   | Leu| 96                | 0.5604 | 96                  | 0.558  | GUG   | Val| 124              | 0.8408 | 123                | 0.8284 |
| CUG   | Leu| 101               | 0.5892 | 99                  | 0.5754 | GUU   | Val| 187              | 1.2676 | 191                | 1.286  |
| CUU   | Leu| 213               | 1.2432 | 212                 | 1.2324 | UGG   | Trp| 153              | 1      | 152                | 1      |
| UUA   | Leu| 248               | 1.4472 | 253                 | 1.4712 | UAC   | Tyr| 75               | 0.5154 | 75                 | 0.512  |
| UUG   | Leu| 201               | 1.173  | 204                 | 1.1862 | UAU   | Tyr| 216              | 1.4846 | 218                | 1.488  |
| UUG   | Met| 1                 | 0.0078 |                     |       |       |     |                   |       |                    |       |

Repeat sequences analysis
Simple sequence repetitions (SSRs) refers to DNA fragments consisting of short units of sequence repetition of 1–6 base pairs in length (B et al., 2014). In this study, we presented the mt genome of *M. multicaulis* there are 386 SSRs were identified by software misa (1.0) and the proportion of different forms showed that SSRs in monomer and trimer was the most abundant types accounted for 86.12% of the total SSRs. There are only One each for pentamer and hexamer and our study also found 6 complex repeating type (Table 4) this phenomenon could contribution to further evolution and genome analyses.

Tandem repeats, also named satellite DNA, widely found in eukaryotic genomes and some prokaryotes (GAO H, 2005), whereas Scattered repetitive sequences are another type of repetitive sequences that different with tandem repetitive sequences, distributed in a dispersed manner in the genome. Use vmatch v2.3.0 software to identify as follows: forward, palindromic, reverse, complement. It is shown that the 30–40 bp repeats are most abundant for both species. In *M. multicaulis* mt genome there were 53 Scattered repetitive sequences and the longest repeat was 22,003 bp were longer than *M. atropurpurea* for 1072bp (Figs. 5 and 6).

**The prediction of RNA editing**

RNA editing that refers to the addition, loss and conversion of the exist in the coding region of the transcribed RNA (Brennicke, Marchfelder, & Binder, 1999) found in all eukaryotes and plants (Malek, Lättig, Hiesel, Brennicke, & Knoop, 1996) that, the conversion of specific cytosine into uridine can alters the genomic information has been reported (§ & Knoop, 2016). We use online sites to predictions (http://www.prepact.de/prepact-main.php), showed that total 377 RNA editing sites within 22 protein-coding genes were predicted while *mttB*, *nad5* and *ccmC* have the most editing sites predicted (32) *atp9* even creat start condon, and remaining 8 protein-coding gene (*atp1*, *atp6*, *atp8*, *cox1*, *cox2*, *cox3*, *rpl16*, *rps19*) does not have any editing site predicted in the mt genome of *M. multicaulis* (Fig. 7). Among those editing sites, 63.93% (241) were located at the second base of the triplet pos. and 36.07% (136) were occurred with the first position of the triplet pos.

In this analysis, the hydrophobicity of 42.18% of amino acids did not change. However, 9.28% of the amino acids were predicted to change from hydrophobic to hydrophilic and 48.54% were predicted to change from hydrophilic to hydrophobic. The RNA editing might lead to the premature termination of protein-coding genes, and this phenomenon showed in Our results that the amino acids of predicted editing codons a leucine tendency after RNA editing (Table 5).
Table 4
Distribution of SSR Type in the mt genomes of *M. multicaulis*.

| Type | SSR | size | NO. | SSR | size | NO. | SSR | size | NO. | SSR | size | NO. |
|------|-----|------|-----|-----|------|-----|-----|------|-----|-----|------|-----|
| p1 (A)10 | 10 | 16 | (C)8 | 8 | 4 | (G)9 | 9 | 1 | (T)14 | 14 | 3 |
| p1 (A)12 | 12 | 2 | (C)9 | 9 | 1 | (T)10 | 10 | 18 | (T)15 | 15 | 1 |
| p1 (A)8 | 8 | 28 | (G)10 | 10 | 1 | (T)11 | 11 | 2 | (T)8 | 8 | 26 |
| p1 (A)9 | 9 | 12 | (G)11 | 11 | 1 | (T)12 | 12 | 1 | (T)9 | 9 | 18 |
| p1 (C)11 | 11 | 2 | (G)8 | 8 | 8 |
| p2 (AG)5 | 10 | 4 | (CT)5 | 10 | 2 | (TA)5 | 10 | 1 | (TC)5 | 10 | 2 |
| p2 (AG)6 | 12 | 2 | (CT)6 | 12 | 2 |
| p3 (AAC)3 | 9 | 1 | (ATT)3 | 9 | 3 | (GAA)3 | 9 | 12 | (GTT)4 | 12 | 2 |
| p3 (AAC)4 | 12 | 2 | (CAA)3 | 9 | 3 | (GAC)3 | 9 | 2 | (TAA)3 | 9 | 3 |
| p3 (AAG)3 | 9 | 4 | (CAC)3 | 9 | 4 | (GAG)3 | 9 | 2 | (TAC)3 | 9 | 1 |
| p3 (AAG)4 | 12 | 2 | (CAG)3 | 9 | 1 | (GAT)4 | 12 | 1 | (TAG)3 | 9 | 3 |
| p3 (AAT)3 | 9 | 3 | (CAT)3 | 9 | 2 | (GCA)3 | 9 | 1 | (TAT)3 | 9 | 2 |
| p3 (ACC)3 | 9 | 2 | (CCA)3 | 9 | 1 | (GCC)3 | 9 | 3 | (TCA)3 | 9 | 1 |
| p3 (ACG)3 | 9 | 2 | (CCG)3 | 9 | 2 | (GCT)3 | 9 | 9 | (TCC)3 | 9 | 1 |
| p3 (ACT)3 | 9 | 2 | (CCT)3 | 9 | 1 | (GCT)4 | 12 | 1 | (TCT)3 | 9 | 5 |
| p3 (ACT)4 | 12 | 1 | (CCT)4 | 12 | 1 | (GGA)3 | 9 | 1 | (TGA)3 | 9 | 1 |
| p3 (AGA)3 | 9 | 11 | (CGC)3 | 9 | 1 | (GGC)3 | 9 | 1 | (TGC)3 | 9 | 1 |
| p3 (AGC)3 | 9 | 6 | (CTA)3 | 9 | 4 | (GTT)3 | 9 | 1 | (TGT)3 | 9 | 1 |
| p3 (AGG)3 | 9 | 4 | (CTC)3 | 9 | 3 | (GTA)3 | 9 | 2 | (TTA)3 | 9 | 4 |
| p3 (AGG)5 | 15 | 1 | (CTG)3 | 9 | 3 | (GTC)3 | 9 | 1 | (TTC)3 | 9 | 10 |
| p3 (AGT)3 | 9 | 1 | (CTT)3 | 9 | 1 | (GTA)3 | 9 | 1 | (TTA)3 | 9 | 1 |
| p3 (ATA)3 | 9 | 7 | (CTT)4 | 12 | 1 | (GTT)3 | 9 | 5 | (TTC)5 | 15 | 1 |
| p3 (ATC)3 | 9 | 3 | (TTG)3 | 9 | 5 |
| p4 (AAAG)3 | 12 | 1 | (CAAA)3 | 12 | 1 | (GAAA)3 | 12 | 1 | (TGCT)3 | 12 | 1 |
| p4 (AACA)3 | 12 | 1 | (CCTT)3 | 12 | 1 | (GTCA)3 | 12 | 1 | (TTAT)3 | 12 | 1 |
| p4 (AAGT)3 | 12 | 1 | (CTTT)3 | 12 | 1 | (TAAG)3 | 12 | 1 | (TTCT)3 | 12 | 3 |
| p4 (AATG)3 | 12 | 1 | (GAAA)3 | 12 | 1 | (GAAA)3 | 12 | 1 | (TGCT)3 | 12 | 1 |
| p4 (ACTC)3 | 12 | 1 | (GACC)3 | 12 | 1 | (TACC)3 | 12 | 1 | (TTTA)3 | 12 | 1 |
| p4 (ATAA)3 | 12 | 1 | (GCCG)3 | 12 | 1 | (GCTT)3 | 12 | 1 | (TTCA)3 | 12 | 1 |
| p4 (ATCT)3 | 12 | 1 | (GCTT)3 | 12 | 1 | (TTCA)3 | 12 | 1 | (TTCT)3 | 12 | 2 |
| p5 (TGAGT)3 | 15 | 1 | (TGAGT)3 | 15 | 1 |
| p6 (AAGGAG)3 | 18 | 1 | (GAAAAG)3 | 18 | 1 |
| c (A)11c(A)8 | 20 | 1 | (T)9(G)9 | 18 | 1 |
| c* (CTA)3(TA)5* | 17 | 1 |
### Table 5
Prediction of RNA editing sites of *M.multicaulis* mt genome

| Type                | Codon   | Aa change | Number | Percentage |
|---------------------|---------|-----------|--------|------------|
| hydrophobic         | TTT->CTT| F->L      | 4      | 28.12%     |
|                     | TTG->CTG| F->L      | 3      | 18.46%     |
|                     | GCT->GTT| A->V      | 1      | 5.26%      |
|                     | GCG->GTG| A->V      | 2      | 10.10%     |
|                     | GCA->GTA| A->V      | 1      | 5.26%      |
|                     | CTT->TTT| L->F      | 11     | 61.05%     |
|                     | CTC->TTC| L->F      | 3      | 15.79%     |
|                     | CCT->CTT| P->L      | 20     | 103.95%    |
|                     | CCG->CTG| P->L      | 19     | 95.45%     |
|                     | CCC->CTC| P->L      | 8      | 40.90%     |
|                     | CTC->CCC| L->P      | 1      | 5.26%      |
|                     | CCA->CTA| P->L      | 33     | 166.76%    |
| hydrophilic         | CGT->TGT| R->C      | 23     | 126.12%    |
|                     | CGC->TGC| R->C      | 9      | 45.45%     |
|                     | CAT->TAT| H->Y      | 13     | 69.09%     |
|                     | CAC->TAC| H->Y      | 8      | 40.90%     |
| hydrophobic-hydrophilic | CCT->TCT| P->S   | 17     | 89.47%     |
|                     | CCG->TCG| P->S      | 5      | 25.26%     |
|                     | CCC->TCC| P->S      | 11     | 55.11%     |
|                     | CCA->TCA| P->S      | 2      | 10.10%     |
| hydrophilic-hydrophobic | TCT->TTT| S->F   | 34     | 173.07%    |
|                     | TCG->TTG| S->L      | 62     | 311.59%    |
|                     | TCA->TTA| S->L      | 49     | 245.26%    |
|                     | ACT->ATT| T->I      | 4      | 20.20%     |
|                     | ACG->ATG| T->M      | 3      | 15.79%     |
|                     | ACC->ATC| T->I      | 1      | 5.26%      |
|                     | ACA->ATA| T->I      | 3      | 15.79%     |
|                     | TCC->CCC| S->P      | 1      | 5.26%      |
|                     | TCA->CCA| S->P      | 2      | 10.10%     |
|                     | CGG->TGG| R->W      | 24     | 120.12%    |

**Homology analysis of chloroplast with mitochondria**

DNA migration is common in plants (Chang et al., 2013). Homologous sequence between chloroplast and mitochondria found using blast software, set similarity to 70%, and E-value to 10E-5 use circos v0.69-5 to visualize it. Twenty-five fragments with a total length of 28,207bp were observed to be migrated from cp genome to mt genome in *M.multicaulis*, accounting for 7.80% of the mt genome (Fig. 8). There are 7 annotated genes located on those fragments, all of which are tRNA genes, namely *trnL-CAA, trnN-GTT, trnM-CAT, trnP-TGG, trnW-CCA, trnD-GTC*, and *tmM-CAT*. Our data also demonstrate that some chloroplast protein-coding genes migrated from cp to mitochondrion,
most of them lost their integrities during evolution, and only partial sequences of those genes could be found in the mt genome such nad1, ccmC, rrn18.

Table 6
Fragments transferred from chloroplast to mitochondria of \textit{M. multicaulis} mt genome

| length | identity | Mis match | Gap opens | mt start | mt end | cp start | cp end | Gene |
|--------|----------|-----------|-----------|----------|--------|----------|--------|------|
| 1      | 3112     | 98.747    | 13        | 3        | 91,640 | 88,555   | 150,948 | 154,059 |
| 2      | 3112     | 98.747    | 13        | 3        | 88,555 | 91,640   | 93,119  | 96,230  |
| 3      | 2936     | 99.251    | 8         | 1        | 100,235| 97,314   | 96,350  | 99,285  |
| 4      | 2936     | 99.251    | 8         | 1        | 97,314 | 100,235  | 147,893 | 150,828 |
| 5      | 2681     | 99.925    | 1         | 1        | 118,160| 115,481  | 134,823 | 137,503 |
| 6      | 2681     | 99.925    | 1         | 1        | 115,481| 118,160  | 112,355 | trnL-CAA |
| 7      | 2180     | 99.083    | 11        | 1        | 346,348| 344,178  | 91,029  | 93,208  |
| 8      | 2180     | 99.083    | 11        | 1        | 344,178| 346,348  | 153,970 | 156,149 |
| 9      | 1073     | 98.788    | 5         | 1        | 349,695| 348,631  | 88,408  | 89,480  |
| 10     | 1073     | 98.788    | 5         | 1        | 348,631| 349,695  | 157,698 | 158,770 |
| 11     | 521      | 87.716    | 18        | 9        | 7,055  | 7,529    | 796    | 1,316   |
| 12     | 235      | 100       | 0         | 0        | 221,648| 221,414  | 89,864  | 90,098  |
| 13     | 235      | 100       | 0         | 0        | 221,414| 221,648  | 157,080 | 157,314 |
| 14     | 889      | 74.241    | 174       | 42       | 152,321| 151,463  | 141,980 | 142,843 |
| 15     | 889      | 74.241    | 174       | 42       | 151,463| 152,321  | 104,335 | 105,198 |
| 16     | 507      | 79.29     | 67        | 25       | 87,604 | 88,091   | 69,809  | 70,296  |
| 17     | 166      | 100       | 0         | 0        | 84,192 | 84,027   | 104,835 | 105,000 |
| 18     | 166      | 100       | 0         | 0        | 84,027 | 84,192   | 142,178 | 142,343 |
| 19     | 156      | 92.308    | 7         | 2        | 122,916| 123,067  | 59,360  | 59,514  |
| 20     | 148      | 92.568    | 10        | 1        | 245,879| 245,732  | 36,734  | 36,880  |
| 21     | 82       | 97.561    | 1         | 1        | 39,222 | 39,142   | 32,340  | 32,421  |
| 22     | 79       | 94.937    | 4         | 0        | 141,020| 140,942  | 55,104  | 55,182  |
| 23     | 62       | 90.323    | 4         | 2        | 164,557| 164,498  | 146,669 | 146,730 |
| 24     | 62       | 90.323    | 4         | 2        | 164,498| 164,557  | 100,448 | 100,509 |
| 25     | 46       | 95.652    | 0         | 1        | 326,639| 326,596  | 45,875  | 45,920  |
| Total  |          |           |           |          |        |          |        | 28,207  |

The different destinations of transferred protein-coding genes and tRNA genes suggested that tRNA genes are much more conserved in the mt genome than the protein-coding genes, indicating their indispensable roles in mitochondria.

Comparison with others green plant mt genomes

Comparison of the mulberry mt genome sequence with other plastomes at the global level using mVISTA online software in shuffle-LAGAN mode of \textit{Morus} species with four family (\textit{Leguminosae, Gramineae, Rosaceae, Asteraceae}). \textit{M. notabilis} used as the reference in the comparative analysis. Interestingly, four family remarkably group-specific and each group shows nearly identical patterns among
themselves. The vista plot patterns produced are remarkably group-specific and each group shows nearly identical patterns among themselves (Fig. 9).

**Variation architecture at the mt genome level**

Nucleic acid diversity (pi) can reveal the variation of nucleic acid sequences of different species, and regions with higher variability can provide potential molecular markers for population genetics. Use maft software (–auto mode) to compare the homologous gene sequences of different species globally, and use dnasp5 to calculate the pi value of each gene. In the mt genome the nucleotide diversity (pi) of the mt genome in cultivated species *M. multicaulis* with wild species of *M. notabilis* and was calculated. In our research found 10 gene (cox1, ccmFc, cob, ccmFN, nad9, mttB, nad3, nad4, atp4, atp9, rps3) pi ranged from 0.00063 to 0.02182 slide window among whole mt genome (Fig. 10). Most of the pi values were lower than 0.01, while rps3 accounting for highest with 0.02182. Besides, total, 85 variations including 79 SNPs and 6 indels were identified across the mt genomes of *M. multicaulis* and *M. atropurpurea* (Table 7). This phenomenon could be applied to further analyses of *Morus* mt genome evolution.

| Summary | Type | Total variation |
|---------|------|-----------------|
|         | SNPs | 79              |
| Indels  | 6    |
| Total   | 85   |

**Phylogenetic analysis within Dicotyledon mt genomes**

To understand the evolutionary status of *Morus* we use MEGA (7.0) to analysis moraceae together with others 7 dicotyledon total 28 species based on the complete mt genome sequence and construct the phylogenetic tree through the ML and NJ methods with a bootstrap of 1,000 replicates to assess the reliability. In this study we collect 28 eudicot from 8 families (*Moraceae*, *Leguminosae*, *Gramineae*, *Brassicaceae*, *Malvaceae*, *Cucurbitaceae*, *Asteraceae*, *Solanaceae*) were well clustered and showed that the phylogenetic tree strongly supports the order of taxa in the phylogenetic tree was consistent with the evolutionary relationships of those species, indicating the consistency of traditional taxonomy with the molecular classification. Based on the phylogenetic relationships among the 28 species, different groups of plants can applied to further comparative analysis. For Moraceae two methods ML and NJ all showed grouped *M. atropurpurea* and *M. multicaulis* together. Which revealed that *M. atropurpurea* and *M. multicaulis* are more related to their congeners than to others and this analysis are important for the mt genome project, development of molecular markers for *Morus* species (Fig. 11 and 12).

**Discussion**

Mitochondria are the power source of the plants that produce the required energy to carry out life processes account of extensive size variations, sequence arrangements, repeat content and highly conserved coding sequence so possess more complex than animals (Kozik et al., 2019). We studied the characteristics of the mt genome of mulberry, a crucial salt tolerance and economically plant with great value as a food and medicine. It is reported that most of the mt genomes is circular, and few are linear such as the mt genome of Polytomella parva in plants (Notsu et al., 2002; Smith, Lee, & Evolution, 2008). In this report, the mt genome of *M. multicaulis* and *M. atropurpurea* has shown that is circular and respectively with 361,546bp and 395,412bp in size, and GC content of the mt genome *Morus* also supports the conclusion that GC content is highly conserved in higher plants.

The repeat sequences contain tandem, short, and large repeats are widely exist in the mt genome (Guo, Zhu, Fan, & Mower, 2016), that play a vital role in shaping the mt genome accounting for those repeats in mitochondria are pivotal for intermolecular recombination (Dong et al., 2018). In this study, we focus on reported the SSRs and scattered repetitive sequences intensively. Research has shown that *M. multicaulis* harbors abundant repeat sequences that might indicate that the intermolecular recombination frequently happens in the mt genome, which maybe applied to dynamically changes the sequence and conformation during the evolution.
RNA-editing is a mean posttranscriptional process that occurs in the both cpDNA and mt genomes of higher plants, which contributing to the better folding of proteins (Bi et al., 2016). So identification of RNA-editing sites provides essential clues for future analysis of predicting gene functions with novel codons about evolution, that can helps us better to understand the gene expression of the cpDNA and mt genomes in plants. Previous studies have been shown that Arabidopsis total 441 RNA-editing sites within 36 genes (Unseld, Marienfeld, Brandt, & Brennicke, 1997), rice have 491 RNA-editing sites within 34 genes (Notsu et al., 2002) and 216 RNA-editing sites within 26 genes of S. glauca (Cheng et al., 2021). Our results show 377 RNA-editing sites within 22 protein-coding genes were predicted of M. multicaulis.

The tRNA genes are much more conserved in the mt genome than the protein-coding genes, indicating their indispensable roles in mitochondria. As the cytoplasmic genome, migration of cpDNA to the mt genome occurred during the plant evolution. We found that 25 fragments were transferred from the cp genome to mt with 7 integrated genes, which are all tRNA genes (Table 6). Transfer of tRNA genes from cp to mt is common in angiosperms (Bi et al., 2016).

We also investigated the mulberry plastome sequence with other land plants at the global level. Conclusively, the genome structure and organization of Morus were consistent together and have a significant differences with other terrestrial green plants. Nucleotide diversity (pi) of the mt genome in M. multicaulis was calculated. In our research found 10 gene ranged from 0.00063 to 0.02182 maybe relative to the evolution. Further, we have analyzed the phylogenetic relationship of mulberry with eudicot representative taxa based on the complete mt genome sequence. Interestingly, Two cultivated are more related than others which familiar to previous reported (QL, Guo, JZ, Yan, & RES, 2016).

Conclusion

In this study, we collected two cultivated species of Morus L. (M. atropurpurea and M. multicaulis) assembled and annotated the mt genome and performed extensive analyses based on the complete mt genome sequences and amino acid sequences of the annotated genes. The Morus L mt genome is circular, M. multicaulis with a length of 361,546 bp. 54 genes, including 31 protein-coding genes, 20 tRNA genes, and 3 rRNA genes, M. atropurpurea is longer than that of M. multicaulis (395,412 bp), total 57 gene contain 32 protein-coding genes, 22 tRNA and 3 rRNA were annotated in the genome.

Our result indicates consistency in molecular and taxonomic classification, besides GC contents in angiosperms, variation and evolutionary status of Morus. This study can provides extensive information about the mt genome for Morus L.

Declarations

Acknowledgments

We thank National Mulberry GeneBank Zhenjiang for providing the original leafs of M. multicaulis and M. atropurpurea.

Authors’ contributions

Guo Liangliang and Zhao Weiguo conceived and designed the research. GL performed, assembled the genomes, analyzed the data, and wrote the original manuscript. Shi Yishu Collected leaf samples, Wu Mengmeng extracted chloroplast DNA., Michael Ackah revised the manuscript, Guo Peng analyze portions of data, Zheng Danyan edited the manuscript, LQ editing of the final manuscript; All authors contributed to the editing of the final manuscript.

Funding

Funding was provided from five department as follows: Sericulture Industry Technology in China Agriculture Research System, Guangxi innovation driven development project, Open Program of Key Laboratory of Silkworm and Mulberry Genetic Improvement Ministry of Agriculture, China, The Crop Germplasm Resources Protection Project of the Agriculture Ministry, and National Infrastructure for Crop Germplasm Resources.

Availability of data and materials

The sequence and annotation of M. multicaulis and M. atropurpurea mt genome data was deposited in the NCBI. The accession number in Gene Banks is MW924382 and MW924383.

Ethics approval and consent to participate
Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

1 Key Laboratory of Silkworm and Mulberry Genetic Improvement, Ministry of Agriculture, School of Biology and Technology, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu, P.R. China, 2Sericultural Research Institute, Guangxi Zhuang Autonomous Regin, Nanning,530007, P.R. China

References

1. §, M. S.-R., & Knoop, V. J. A. i. B. R. (2016). Coevolution of Organelle RNA Editing and Nuclear Specificity Factors in Early Land Plants - ScienceDirect. 78, 37-93. 10.1016/bs.abr.2016.01.002
2. Alverson, A. J., Xiaoxin, W., Rice, D. W., Stern, D. B., Kerrie, B., Palmer, J. D. J. M. B., & Evolution. (2010). Insights into the Evolution of Mitochondrial Genome Size from Complete Sequences of Citrullus lanatus and Cucurbita pepo (Cucurbitaceae). (6), 1436. 10.1093/molbev/msg029
3. Archibald, J. M. J. S. (2011). Origin of eukaryotic cells: 40years on. 54(2), 69-86. 1007/s13199-011-0129-z
4. B, Y. C. L. A., A, S. L., C, D. C. L., A, Y. X. W., A, C. L., A, Y. M. Y., . . . A, W. S. L. J. S. H. (2014). Exploiting EST databases for the development and characterization of EST-SSR markers in blueberry (Vaccinium) and their cross-species transferability in Vaccinium spp - ScienceDirect. 176, 319-329. 10.1016/j.scienta.2014.07.026
5. Bi, C., Paterson, A. H., Wang, X., Xu, Y., Wu, D., Qu, Y., . . . Ye, N. J. B. R. I., . (2016). Analysis of the Complete Mitochondrial Genome Sequence of the Diploid Cotton Gossypium raimondii by Comparative Genomics Approaches. 2016, 1-18. 10.1155/2019/9691253
6. Brennicke, A., Marchfelder, A., & Binder, S. J. F. M. R. (1999). RNA editing. 23(3), 297-316. 10.1111/j.1574-6976.1999.tb00401.x
7. Chan, W. C., Lye, P. Y., Wong, S. K., Sciences, F., University, U., Science, S. O., & Medicines, M. U. S. J. C. J. o. N. (2016). Phytochemistry,pharmacology,and clinical trials of Morus alba. 10.3724/SP.J.1009.2016.00017
8. Chang, S., Wang, Y., Lu, J., Gai, J., Li, J., Chu, P., . . . Zhao, T. J. P. O. (2013). The Mitochondrial Genome of Soybean Reveals Complex Genome Structures and Gene Evolution at Intercellular and Phylogenetic Levels. 8. 10.1371/annotation/5bf22546-6983-42c9-9cb5-1a6459b29a79
9. Cheng, Y., He, X., Priyadarshani, S. V. G. N., Wang, Y., & Qin, Y. J. B. G. (2021). Assembly and comparative analysis of the complete mitochondrial genome of Suaeda glauca. 22(1). 10.1186/s12864-021-07490-9
10. Daniel, B., Sloan, Andrew, J., Alverson, . . . Biology, M. J. P. (2012). Rapid Evolution of Enormous, Multichromosomal Genomes in Flowering Plant Mitochondria with Exceptionally High Mutation Rates. 70(1), e1001241. 10.1371/journal.pbio.1001241
11. Dong, S., Zhao, C., Chen, F., Liu, Y., Zhang, S., Wu, H., . . . Yang, L. J. B. G. (2018). The complete mitochondrial genome of the early flowering plant Nymphaea colorata is highly repetitive with low recombination. 79(1), 1-12. 10.1186/s12864-018-4991-4. 1186/s12864-018-4991-4
12. Drouin, G., Daoud, H., Xia, J. J. M. P., & Evolution. (2008). Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. 49(3), 827-831. 10.1016/j.ympev.2008.09.009
13. GAO H, K. J. (2005). Distribution Characteristics and Biological Function of Tandem Repeat Sequences in the Genomes of Different Organisms. 026(005), 555-564. 10.1360/crad20050909
14. Guo, W., Zhu, A., Fan, W., & Mower, J. P. J. N. P. (2016). Complete mitochondrial genomes from the ferns Ophioglossum californicum and Psilotum nudum are highly repetitive with the largest organellar introns. 273. 1111/nph.14135
15. He N, Z. C., Qi X, Zhao S, Tao Y, Yang G, Lee TH, Wang X, Cai Q, Li D, et al. (2013). Draft genome sequence of the mulberry tree Morus notabilis %J Nature Communications. 4. 10.1007/s11105-015-0877-7
16. Huang, C. Y., Ayliffe, M. A., & Timmis, J. N. J. N. (2003). Direct measurement of the transfer rate of chloroplast DNA into the nucleus. 422(6927), 72-76. 10.1038/nature01435
17. Kozik, A., Rowan, B. A., Lavelle, D., Berke, L., & Christensen, A. C. J. P. G. (2019). The alternative reality of plant mitochondrial DNA: One ring does not rule them all. 15(8), e1008373-. 10.1371/journal.pgen.1008373

18. Malek, O., Lättig, K., Hiesel, R., Brennicke, A., & Knoop, V. J. E. J. (1996). RNA editing in bryophytes and a molecular phylogeny of land plants. 15(6), 1403-1411. 10.1002/j.1460-2075.1996.tb00482.x

19. Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., . . . Genomics. (2002). The complete sequence of the rice (Oryza sativa L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. 268(4), 434-445. 10.1002/s00438-002-0767-1

20. QL, Guo, JZ, Yan, & RES, C. J. G. M. (2016). Complete chloroplast genome sequence of cultivated Morus L. species. 2016,15(4), -. 10.4238/gmr15048906

21. Rehman, J., Zhang, H. J., Toth, P. T., Zhang, Y., Marsboom, G., Hong, Z., . . . Archer, S. L. J. F. J. O. P. o. t. F. o. A. S. f. E. B. (2012). Inhibition of mitochondrial fission prevents cell cycle progression in lung cancer. 26(5), 2175-2186. 10.1096/fj.11-196543

22. Richardson, A. O., Rice, D. W., Young, G. J., Alverson, A. J., & Palmer, J. D. J. B. B., 11,1. (2013). The "fossilized" mitochondrial genome of Liriodendron tulipifera: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. 11. 10.1186/1741-7007-11-29

23. Sebastian, Beier, Thomas, Thiel, Münch, Uwe, . . . Bioinformatics, M. J. (2017). MISA-web: a web server for microsatellite prediction. 10.1093/bioinformatics/btx198

24. Sergey, Koren, Brian, P., Walenz, Konstantin, Berlin, . . . research, B. J. G. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. 10.1101/gr.215087.116

25. Shtolz, N., Dan, M. J. F. i. E., & Evolution. (2019). The Mitochondrial Genome—on Selective Constraints and Signatures at the Organism, Cell, and Single Mitochondrion Levels. 7. 10.1111/jeb.12822

26. Skippington, E., Barkman, T. J., Rice, D. W., & Palmer, J. D. J. P. N. A. S. U. S. A. (2015). Miniaturized mitogenome of the parasitic plant Viscum scurruloideum is extremely divergent and dynamic and has lost all nad genes. 112(27), E3515. 10.1073/pnas.1504491112

27. Smith, D. R., Lee, R. W. J. M. B., & Evolution. (2008). Mitochondrial Genome of the Colorless Green Alga Polytomella capuana: A Linear Molecule with an Unprecedented GC Content. 25(3), 487-496. 10.1093/molbev/msm245

28. Tong, W., Kim, T. S., & Park, Y. J. J. R. (2016). Rice Chloroplast Genome Variation Architecture and Phylogenetic Dissection in Diverse Oryza Species Assessed by Whole-Genome Resequencing. 9(1). 10.1038/srep43327

29. Unseld, M., Marienfeld, J. R., Brandt, P., & Brennicke, A. J. N. G. (1997). The mitochondrial genome of Arabidopsis thaliana contains 57 genes in 366,924 nucleotides. 15(1), 57-61. 10.1007/s00122-004-1895-2

30. Walker, B. (2015). Pilon: Comprehensive Microbial Variant Detection and Genome Assembly Improvement. 10.1371/journal.pone.0112963

31. Wolfe, K. H., Li, W. H., & Sharp, P. M. J. P. o. t. N. A. o. S. (1988). Wolfe KH, Li WH, Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNA. Proc Natl Acad Sci USA 84: 9054-9058. 84(24), 9054-9058. 10.1073/pnas.84.24.9054

32. Woloszynska, M. J. J. o. E. B. (2009). Heteroplasmy and stoichiometric complexity of plant mitochondrial genomes—though this be madness, yet there's method in’t. 10.1093/jxb/erp361

33. X., & Physiology, T. J. P. (2006). The Rice Mitochondrial Genomes and Their Variations. 10.2307/4282065

34. Xiao-fang, Liao, Yanhong, Zhao, Xiang-jun, Kong, . . . Reports, Z. J. S. (2018). Complete sequence of kenaf (Hibiscus cannabinus) mitochondrial genome and comparative analysis with the mitochondrial genomes of other plants. 10.1038/s41598-018-30297-w

35. Yasunari, O., Yukiko, Y., Koji, M., Akira, K., Toru, T., Takashi, S., . . . Naoki, M. J. N. A. R. (2005). Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. (19), 6235. 10.1093/nar/gki925

Figures
Figure 1

The circular map of the mt genome of *M.multicaulis*

Figure 2

The circular map of the mt genome of *M.atropurpurea Roxb*
The circular map of the mt genome of *M. atropurpurea*. The forward direction gene encoding is located outside the circle, and reverse direction is located inside. The inner gray circle represents the GC content.

**Figure 3**

Relative Synonymous Codon Usage analysis of *M. multicaulis*

**Figure 4**

Relative Synonymous Codon Usage analysis of *M. atropurpurea*
**Figure 5**

Scattered repetitive sequence of *M. multicaulis*

**Figure 6**

Scattered repetitive sequence of *M. atropurpurea*
Figure 7

The distribution of RNA-editing sites in M.multicaulis mt genome protein-coding.

Figure 8

DNA migration from chloroplast to mitochondria of M.multicaulis.
Figure 9

Percent identity plot for comparison of three Morus L relative to Eudiots.

Figure 10
The nucleotide diversity (pi) of M.multicaulis mt genome

Figure 11

Phylogenetic analysis of Morus species using the complete mt genome by the ML method.
Figure 12

Phylogenetic analysis of Morus species using the complete mt genome by the NJ method.