The Complete Chloroplast Genome Sequence and Phylogenetic Analysis of Mulberry (*Morus alba* Lin.)

Guo Liangliang¹, Shi Yisu¹, Wu Mengmeng¹, Michael Ackah¹, Qiang Lin² and Weiguo Zhao^{1*}

¹Key Laboratory of Silkworm and Mulberry Genetic Improvement, Ministry of Agriculture, School of Biology and Technology, Jiangsu University of Science and Technology, P.R. China

²Sericultural Research Institute, Guangxi Zhuang Autonomous Region, China

*Corresponding author: Weiguo Zhao, 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

ARTICLE INFO

Received: November 16, 2021

Published: November 24, 2021

Citation: Guo Liangliang, Shi Yisu, Wu Mengmeng, Michael Ackah, Qiang Lin, Weiguo Zhao. The Complete Chloroplast Genome Sequence and Phylogenetic Analysis of Mulberry (*Morus alba* Lin.). Biomed J Sci & Tech Res 40(2)-2021. BJSTR. MS.ID.006418.

Keywords: *M. multicaulis*; *M. atropurpurea*; Complete Chloroplast Genome; Phylogenetic

ABSTRACT

The chloroplast (cp) is the photosynthetic organelle representing one of the most important organelles in green plants and algae, with strong saline-alkali tolerance. It is also admired as a landscape plant with high development prospects and scientific research value. In this study, The cp genome of two cultivated species were assembled based on the reads from Pacbio and Illumina sequencing platforms. The full length of *M. multicaulis* chloroplast genome is 158,817 bp, including a pair of inverted repeats (IR) region of 25,551 bp, a large single-copy (LSC) region of 87,880 bp and a small single-copy (SSC) region of 19,835bp long. *M. atropurpurea* consists of an LSC region (87,670), an SSC region (19,750 bp) and (IRa and IRb) are both 25,667 bp which is shorter than *M. multicaulis* (158,776bp). Each cpDNA contains 113 functional genes: including 79 protein-coding genes, 30 tRNA genes and 4 rRNA genes. The GC content of *M. multicaulis* and *M. atropurpurea* were 36.24% and 36.26% respectively. The MEGA-X was used to construct a phylogenetic tree 26 species. The result shows that *M. atropurpurea* and *M. multicaulis* more related to their congeners than to others. This study provides more detailed information on cpDNA evolution and its structural analysis are important for the chloroplast genome project, development of molecular markers for *Morus* species.

Introduction

The chloroplast(cp) is the photosynthetic organelle representing one of the most important organelles in green plants and algae [1]. The origin of chloroplasts can be dated back to about 390,000 years ago [2]. The mainstream view is the theory of endosymbiosis, which believes that plant chloroplasts originated from ancient cyanobacteria that were symbiotic in primitive eukaryotic cells [3]. In angiosperms, the chloroplast genome (cpDNA) is typically composed of a pair of inverted repeat regions (IRa and IRb), which are separated by a small single-copy (SSC) region and a large single-copy (LSC) region [4]. The contraction and expansion of IR region [5,6] determine the length of the chloroplast genome.

Most cp genomes are 120–220kb in length and contain 120–140 coding genes (Marcelo, Leila, Fraga, & Guerra, 2015; Méndez-

Leyva et al). The size of chloroplasts are similar in related species [7]. The GC content in the chloroplast genome is generally low, about 37% on average, with AT tendency [8]. The chloroplast sequence is widely used to phylogenetics, species identification, population genetics, and genetic engineering [9-11]. Recently, reports on phylogenetic analysis in rice [12], analysis clarify the taxonomic status of Capsicum L (Elmosallamy et al. 2019) and Korean ginseng [13]. The cpDNA sequence of Moraceae is incomplete and limited. The results of phylogenetic study on mulberry based on nuclear genome redefined the mulberry as eight species [14]. But the conclusions are not enough to reveal the complex systemic origin, evolutionary relationship, and phylogenetic studies of mulberry. DNA molecular markers like indels, SSRs and small inversions [15], ITS regions [16,17] have been used to study genetic and genome

diversity and phylogenetic analysis. In this study, the cpDNA sequences of *M. atropurpurea* and *M. multicaulis* were investigated, and a comparative analysis was performed between cultivated *Morus* and *M. mongolica*. The genome structure, gene order, repeat sequences, and phylogenetics were analyzed.

Materials and Methods

Plant material, DNA Extraction, and Sequencing

M. atropurpurea (Lunjiao40) and *M. multicaulis* (Husang32) fresh leaf were collected from National Mulberry Genebank Zhenjiang City, Jiangsu Province. The plants DNA were extracted using CTAB method for complete and high purity DNA to build DNA library following manufacture's protocol. The DNA concentration was measured by Nanodrop instrument then qualified samples were sent to the Illumina NovaSeq for sequencing.

Quality Control of Sequencing Data

To improve the accuracy of the analysis, the Raw Reads were filtered again according to the following criteria:

1. Removal of the sequenced connectors and primer sequences in reads,
2. Reads with an average mass value less than Q5 were filtered out,
3. Reads with N number greater than 5 were removed. The quality reads after the above checks, called clean reads were subjected to subsequent analysis.

Assembly and Annotation

The SPAdes (3.13.0) software was used for the genome splicing [18], The candidate sequence assembly was determined, annotate, and then GeSeq was used to draw the circular gene map.

Analysis of Repeated Sequences

Long repeats include three types: forward (P), palindrome (P)

and tandem (T) repeats, which may promote chloroplast genome rearrangement and increase population genetic diversity. We use vmatch (<http://www.vmatch.de/>) The software (parameter: minimal repeat size 30bp) finds the scattered long repeat fragments in the chloroplast genome.

Comparative Analysis of Chloroplast Genomes of Morus Species with Other Species

The mVISTA online software in shuffle-LAGAN mode was applied to compare the complete chloroplast genomes of cultivated *Morus* species with four representatives. *M. multicaulis* (husang32) was used as a reference. Also, different families *Arabidopsis thaliana* (NC000923.1), *Fragaria chiloensis* (JN884816), *Oryza sativa* (NC-008155), *Nicotiana tabacum* (Z00044) were analysed together. The software (parameter: minimal repeat size 30bp) finds the scattered long repeat fragments in the chloroplast genome.

Phylogenetic Analysis

The MEGA X software was used to determine the phylogenetic relationships between *Morus* species by the maximum likelihood (ML) and neighbor-joining (NJ) methods. Data on the cpDNA of *Morus* species are available, including those from including those from NCBI *M. indica* (NC-008359), *M. mongolica* (KM491711), *M. notabilis* (KP939360).

Results

Analysis of Sequencing Data and Quality Control

An overview of the chloroplast sequencing reads derived from the *M. multicaulis* *M. atropurpurea* libraries is listed in Table1. A total of 30417184 and 32296130 raw reads were obtained from *M. multicaulis* and *M. atropurpurea*, respectively. After quality control check on the raw reads, 30374110 and 32246870 clean reads were obtained from *M. multicaulis*, and *M. atropurpurea* respectively. In addition, the Median Phred Quality Score as well as the Read Cycle in mulberry plants are shown in Figure 1.

Figure 1: Rawdata of *M. multicaulis* and *M. atropurpurea*.

Table 1: Rawdata of Morus.

Material_	length(bp)	Raw_Reads	Raw_Bases(bp)	Clean_Reads	Clean_Bases(bp)	GC(%)	Q20(%)	Q30(%)
<i>M. multicaulis</i>	150	30,417,184	4,562,577,600	30,374,110	4,556,116,500	35.27	96.86	91.8
<i>M. atropurpurea</i>	150	32,296,130	4,844,419,500	32,246,870	4,837,030,500	35.6	97.13	92.4

Table 2: Comparison of chloroplast genomes between seven *Morus L* species.

Chloroplast feature	<i>M.indica</i>	<i>M. notabilis</i>	<i>M. mongolica</i>	<i>M. atropurpurea</i>		<i>M. multicaulis</i>	
				Yichuanhong	Lunjiao40	Ribentiancheng	Husang32
Genome size (bp)	158484	158680	158459	159113	158,776	159154	158,817
Total GC content (%)	36.4	36.4	36.3	36.2	36.26	36.2	36.24
genes (#)	133	129	133	130	113	130	113
protein-coding genes (#)	88	84	88	85	79	85	79
LSC Length (bp)	87386	87470	87367	87824	87,670	87763	87,880
percent (%)	55.14	55.12	55.14	55.19	55.21	55.15	55.33
GC content (%)	34.1	34.1	34	33.9	33.91	33.9	33.9
SSC length (bp)	19742	19776	19736	19875	19,750	20035	19,835
percent (%)	12.46	12.46	12.45	12.49	12.44	12.59	12.49
GC content (%)	29.4	29.3	29.3	29.3	29.34	29.3	29.24
IR length (bp)	25678	25717	25678	25707	25,667	25707	25,551
percent (%)	16.2	16.21	16.2	16.16	16.17	16.16	16.09
GC content (%)	42.9	42.9	42.9	42.9	42.92	42.9	42.99

Table 3: Genos present in the chloroplast genome of *M. atropurpurea* and *M. multicaulis*.

Function	Gene group	Gene_Name
Self-replication	ribosomal RNAs	rrn16, rrn4.5, rrn5, rrn23
	transfer RNAs	trnV-GAC, trnT-UGU, trnM-CAU, trnT-GGU, trnPUGG, trnE-UUC, trnR-ACG, trnH-GUG, trnS-GCU, trnM-CAU, trnLUAA, trnQ-UUG, trnK-UUU, trnD-GUC, trnC-GCA, trnS-GGA, trnF-GAA, trnV-UAC, trnG-GCC, trnA-UGC, trnL-UAG, trnI-GAU, trnR-UCU, trnY-GUA, trnG-UCC, trnN-GUU, trnL-CAA, trnI-CAU, trnW-CCA, trnS-UGA
	ribosomal proteins (LSU)	rpl2, rpl33, rpl36, rpl22, rpl16, rpl14, rpl20, rpl32, rpl23
	ribosomal proteins (SSU)	rps14, rps3, rps15, rps8, rps4, rps12, rps18, rps19, rps7, rps16, rps11, rps2
	RNA polymerase	rpoB, rpoA, rpoC1, rpoC2
	NADH dehydrogenase	ndhA, ndhG, ndhJ, ndhK, ndhB, ndhE, ndhF, ndhH, ndhI, ndhD, ndhC
Photosynthesis	photosystem I	psaJ, psaA, psaC, psaI, psaB
	photosystem II	psbZ, psbB, psbE, psbF, psbC, psbK, psbA, psbH, psbD, psbM, psbT, psbJ, psbL, psbI
	cytochrome b/f complex	petB, petD, petL, petN, petA, petG
	ATP synthase	atpA, atpI, atpB, atpF, atpH, atpE
	RubisCO large subunit	rbcl
	Maturase	matK
	Protease	clpP
	other genes	accD, cemA, pbf1, ccsA
Unknown function	hypothetical chloroplast reading frames(ycf)	ycf1(2), ycf3, ycf2, ycf4

Codon Usage

All 79 protein-coding genes in the cpDNA of *M. multicaulis* and *M. atropurpurea* were encoded by 60,765 codons (Table 4). Codon usages have an AT tendency. For *M. multicaulis* 63.76% of codons ending with A or T. Amino acid Leu has the most codons usage of

6,435 (10.59%) codons, followed by Ile with 5310. Ser with 4,464 and Glu with 4,464. These four amino acid codons represented one-third of the total. The least Cys had 699 codons (1.15%). AUU is the most frequent start codon (872), and UGA is the most common stop codon (11).

Table 4: Codon usage for *M. multicaulis* and *M. atropurpurea*.

Condon	AA	<i>M.multicaulis</i>		<i>M.atropurpurea</i>		Condon	AA	<i>M.multicaulis</i>		<i>M.atropurpurea</i>	
		Number	RSCU	Number	RSCU			Number	RSCU	Number	RSCU
UAA	*	27	1.69	27	1.69	AUG	M	457	1	457	1
UAG	*	10	0.62	10	0.62	AAU	N	811	1.57	811	1.57
UGA	*	11	0.69	11	0.69	AAC	N	225	0.43	225	0.43
GCU	A	503	1.86	503	1.86	CCU	P	317	1.54	317	1.54
GCC	A	163	0.6	162	0.6	CCC	P	146	0.71	146	0.71
GCA	A	301	1.11	301	1.11	CCA	P	239	1.16	239	1.16
GCG	A	116	0.43	116	0.43	CCG	P	120	0.58	120	0.58
UGU	C	177	1.52	177	1.52	CAA	Q	579	1.56	579	1.56
UGC	C	56	0.48	56	0.48	CAG	Q	164	0.44	164	0.44
GAU	D	678	1.61	678	1.61	CGU	R	267	1.35	267	1.35
GAC	D	163	0.39	163	0.39	CGC	R	77	0.39	77	0.39
GAA	E	862	1.54	862	1.54	CGA	R	260	1.32	260	1.32
GAG	E	259	0.46	259	0.46	CGG	R	77	0.39	77	0.39
UUU	F	784	1.37	784	1.37	AGA	R	375	1.9	375	1.9
UUC	F	362	0.63	362	0.63	AGG	R	130	0.66	130	0.66
GGU	G	469	1.36	469	1.36	UCU	S	418	1.69	418	1.69
GGC	G	148	0.43	148	0.43	UCC	S	236	0.95	236	0.95
GGA	G	549	1.59	549	1.59	UCA	S	300	1.21	300	1.21
GGG	G	212	0.62	212	0.62	UCG	S	130	0.52	130	0.52
CAU	H	380	1.54	380	1.54	AGU	S	309	1.25	309	1.25
CAC	H	113	0.46	113	0.46	AGC	S	95	0.38	95	0.38
AUU	I	872	1.48	872	1.48	ACU	T	413	1.67	413	1.67
AUC	I	333	0.56	333	0.56	ACC	T	187	0.76	187	0.76
AUA	I	565	0.96	565	0.96	ACA	T	287	1.16	287	1.16
AAA	K	823	1.52	823	1.52	ACG	T	102	0.41	102	0.41
AAG	K	260	0.48	260	0.48	GUU	V	396	1.47	397	1.47
UUA	L	701	1.96	701	1.96	GUC	V	116	0.43	116	0.43
UUG	L	433	1.21	433	1.21	GUA	V	426	1.58	426	1.57
CUU	L	460	1.29	460	1.29	GUG	V	143	0.53	143	0.53
CUC	L	128	0.36	128	0.36	UGG	W	362	1	362	1
CUA	L	289	0.81	289	0.81	UAU	Y	630	1.62	629	1.62
CUG	L	134	0.37	134	0.37	UAC	Y	147	0.38	148	0.38

Repeat Sequences Analysis

Long repeats include three types: forward (P), palindromic (P) and tandem (T) repeats, which may promote chloroplast genome rearrangement and increase population genetic diversity. The *M.*

atropurpurea with the most extended repetitive sequence located in 133,098 with 25,678 bp, which is 127 bp longer than that of *M. multicaulis* and the shortest is 31 bp, which is 8 bp shorter than that of *M. multicaulis* (Tables 5 & 6). This is a phenomenon of genome rearrangement and may be related to genetic diversity.

Table 5: *M. multicaulis* Long repeat sequence analysis.

Length	Position	Type	Length	Position	Distance_value	E-value	Score_value	Percent_identity
39	45935	F	39	101,936	0	2.35E-14	78	100
25551	87880	P	25551	133,266	0	0.00E+00	51,102	100
39	45935	P	39	144,722	0	2.35E-14	78	100
30	125482	P	30	125,482	0	6.15E-09	60	100
127	0	P	127	87,753	0	2.45E-67	254	100
30	9104	P	30	47,631	0	6.15E-09	60	100

Table 6: *M.atropurpurea* Long repeat sequence analysis.

Length	Position	Type	Length	Position	Distance_value	E-value	Score_value	Percent_identity
39	45866	F	39	101,853	0	2.35E-14	78	100
31	33278	F	31	33,302	0	1.54E-09	62	100
25678	87670	P	25,678	133,098	0	0.00E+00	51,356	100
39	45866	P	39	144,554	0	2.35E-14	78	100
30	125314	P	30	125,314	0	6.15E-09	60	100
30	8984	P	30	47,562	0	6.15E-09	60	100

Comparison with Other Species Chloroplast Genomes

We compared with others 6 species and two exogenous *Pyrus pyrifolia* and *Zea mays* based on the complete cp genomic sequences (Figure 3). The results showed that the IR boundary of *Morus* chloroplast genome is slightly different. The *rp12* gene at the junction of LSC/IRb 67bp in LSC and the rest are in IRb and *trnH* gene which is 150bp away from the IRa/LSC boundary of *M. multicaulis*. *Pyrus pyrifolia* is 101 bp away from the IRa/LSC boundary. The junction IRb/SSC 52bp of *ndhF* gene is located at IRb, and the rest is located at SSC of *M. alba*. *Ycf1* is located in IRa/

SSC boundary, resulting in the formation of *ycf1* pseudogene. The *rpsl9* gene is located at the LSC/IR boundary, which resulted in the formation of a *rpsl9* pseudogene. These results are similar to previous study findings [19]. We analyzed the chloroplast genome sequences of 7 mulberry species and *Arabidopsis thaliana*, *Fragaria chiloensis*, *Oryza sativa*, *Nicotiana tabacum* using mVISTA software (Figure 4). Results showed that the chloroplast genome within the genus *Morus* was relatively similar. The regions with higher variation were distributed in the LSC and SSC regions, and the IR region was more conservative.

Figure 4: Percent identity plot for comparison of seven *Morus L. chloroplast* genomes using m VISTA.

Phylogenetic Analysis

Using MEGA-X (10.2.2) software through the ML (Figure 5) and NJ methods (Figure 6) cluster analysis based on the complete

chloroplast genome sequences. *M. multicaulis* (Husang32) and *M. atropurpurea* (Lunjiao40) are grouped together it's means they are closely related and diverged from the *M. indica* and *M. notabilis* earlier.

Figure 5: Phylogenetic analysis of the *Morus L.* using complete chloroplast genome by the ML method.

Figure 6: Phylogenetic analysis of *Morus* species using the complete chloroplast genome by the NJ method.

Discussion

In this study, we collected two cultivated species of *Morus* L. (*M. atropurpurea* and *M. multicaulis*) assembled and annotated the cp genome and performed extensive analyses based on the complete cp genome sequences and amino acid sequences of the annotated genes. The *Morus* L cp genome is circular contain LSC, IR, SSC. compared with wild species of *M. mongolica*, *M. notabilis* and *M. indica* in the genome sequence, providing more detailed information for phylogenetic studies. The results show that the chloroplast genome length of 7 mulberry species ranges from 158,459 to 159,154 bp. *M. atropurpurea* and *M. multicaulis* with two species different accession also have slight differences. "Husang32" and "Ribentiancheng" of *M. multicaulis*, with "Husang32" 377bp being shorter than the "Ribentiancheng" (Table 1). This result agrees with the gene sequence used by Li Qiaoli [20] on *M. multicaulis*. Also, *M. atropurpurea* "Lunjiao40" is 337bp shorter than "Yichuanhong" (Table 2). There were few differences in the length of the IR and SSC regions of the cpDNA from seven species, but most of the differences mainly focus on the LSC region. The results indicated that these species are closely related and confirmed by the phylogenetic tree analysis.

IR expansion/contraction studies reveal the considerable difference even in the same family [13]. It is believed that the borders of the IR region (IRa and IRb) with the LSC and SSC regions, some expansion and contraction may play an important role in evolution [21]. Therefore, it is of great interest to compare IR/SC junctions in different varieties. Based on this, we use 6 mulberry species along with other members, complete chloroplast genome sequences. The results showed that rp12 gene located in the junction of LSC/IRb 67bp in LSC and the rest are in IRb and trnH gene is 150 bp away from the IRa/LSC boundary of *M. multicaulis*. With ndhF located in the junction of IRb/SSC 52bp at IRb, the rest is located at SSC of *M. alba*. Based on IR boundary, it could be seen that *M. atropurpurea* and *M. multicaulis* are closely related and together also closely related to *M. mongolica* and *M. notabilis*. This result agrees with the finding by Li Qiaoli [20].

Usually, cpDNA also tends to AT [7,22]. The AT content of the four regions of the seven species (Table 2) shown The IR region has the lowest with an average of only 57.08%, LSC region and the SSC region are 66.02% and 70.67%, *M. atropurpurea*, and *M. multicaulis* were slightly higher than other groups. The average of 4 species was about 63.78%, SC regions with a high AT content harbour more variation SSR polymorphisms between *M. multicaulis* and *M. atropurpurea* are all involved A or T mutations. The lowest AT content in the IR region maybe because it contains all the rRNA, which is very conservative, and the GC content is high and relatively stable. The rpl21 gene is known to only exist in the plastomes of ferns and bryophytes [23]. This result shows that *Morus* contains two pseudogenes ycf1 and rps19. SSR as an important molecular

marker used for studying population genetics [24,25]. Which may suggest their usefulness in future evolutionary studies on *Morus* species. With the Phylogenetic analysis, seven species of *Morus* gathered together. Both *M. atropurpurea* and *M. multicaulis* gather together, which is the closest. This result is similar to Li Qiaoli [20]. In the present research, cpDNA sequences of *Morus* species are minimal. Our experimental results do not represent the final classification of the *Morus* system. Therefore, the evolutionary relationship of *Morus* needs further research to draw more accurate conclusions. Reported here enhance genome information on *Morus* and contribute to the study of germplasm diversity. These data represent a valuable Phylogenetic analysis source for future studies on *Morus* populations diversity. Genetic diversity within the chloroplast is considered effective tool for understanding population genetic structure and species evolution. We assembly the full chloroplast DNA sequences of two cultivated species of *M. atropurpurea* and *M. multicaulis* and compared with other six wild *Morus* species. These sequences enabled us to identify evolution divergence time of *Morus* and can be used for further research of genetic diversity, genetic structure, and genome evolution history of *Morus* [26,27].

Acknowledgement

This work was supported by China Agriculture Research System of MOF and MARA (CARS-18-ZJ0207), National Key R&D Program of China, key projects of international scientific and technological innovation cooperation (2021YFE0111100), Guangxi innovation-driven development project (AA19182012-2), Zhenjiang Science and Technology support project (GJ2021015).

Author Contribution Statement

GL: conceived and designed the project, assembled the genomes, analyzed the data, and wrote the original manuscript; SY: designed the experiments, analyze portions of data, revised and edited the manuscript; WM, MA, , LQ: Collected leaf samples and extracted chloroplast DNA. ZW: supervised the experiment and edited the manuscript. All authors contributed to the editing of the final manuscript.

Data Availability Statement

The whole chloroplast genome data have been deposited at NCBI with the following accession numbers: MW548981 and MW548982. The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

MW548981 (<https://www.ncbi.nlm.nih.gov/nucleotide/MW548981>)

MW548982 (<https://www.ncbi.nlm.nih.gov/nucleotide/MW548982>)

Declarations

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Competing Interests

The authors declare that they have no competing interests.

References

- Daniell H, HT Chan, Pasoreck EK (2016) Vaccination via Chloroplast Genetics: Affordable Protein Drugs for the Prevention and Treatment of Inherited or Infectious Human Diseases. *Annu Rev Genet* 50: 595-618.
- Hohmann N, Wolf EM, Rigault P, WB Zhou, Kiefer M, et al. (2018) Ginkgo biloba's footprint of dynamic Pleistocene history dates back only 390,000years ago. *BMC Genomics* 19: 299.
- Howe CJ, Barbrook AC, Koumandou VL, Nisbet RER, Symington HA, et al. (2003) Evolution of the chloroplast genome. *Philos Trans R Soc Lond B Biol Sci* 358: 99-106.
- Daniell H LS-B, Grevich J, Saski C, Quesada-Vargas T, Guda C, et al. (2006) Complete chloroplast genome sequences of *Solanum bulbocastanum*, *Solanum lycopersicum* and comparative analyses with other Solanaceae genomes. *Theor Appl Genet* 112: 1503.
- W Dong, X Chao, C Tao, S Zhou, Nicolas SJ (2013) Complete Chloroplast Genome of *Sedum sarmentosum* and Chloroplast Genome Evolution in Saxifragales. *Plos One* 8: e77965.
- HY Zhang, C Li, HM Miao, SJ Xiong (2013a) Insights from the Complete Chloroplast Genome into the Evolution of *Sesamum indicum* L. *Plos One* 8: e80508.
- JX Liu (2015) De novo Assembly and Gene Annotation of the Chloroplast Genomes of *Lycium barbarum* and *Lycium ruthenicum*. Dissertation, Beijing Institute of Genomics Chinese Academy of Sciences.
- Daniell H, Wurdack KJ, Kanagaraj A, Lee SB, Saski C, et al. (2008) The complete nucleotide sequence of the cassava (*Manihot esculenta*) chloroplast genome and the evolution of *atpF* in Malpighiales: RNA editing and multiple losses of a group II intron. *Theor Appl Genet* 116: 723-737.
- Asaf S, Khan AL, Khan MA, Waqas M, SM Kang, et al. (2017) Chloroplast genomes of *Arabidopsis halleri* ssp. *gemma* and *Arabidopsis lyrata* ssp. *petraea*: Structures and comparative analysis. *Sci Rep* 7: 7556.
- S Xue, T Shi, WJ Luo, XP Ni, Lqbal S, et al. (2019) Comparative analysis of the complete chloroplast genome among *Prunus mume*, *P. armeniaca*, and *P. salicina*. *Hortic Res* 6: 89.
- YD Gao, Harris A, XJ He (2015) Morphological and ecological divergence of *Lilium* and *Nomocharis* within the Hengduan Mountains and Qinghai-Tibetan Plateau may result from habitat specialization and hybridization. *BMC Evol Biol* 15: 147-168.
- L Cheng, Nam J, SH Chu, Runghana P, Park YJ (2019) Signatures of differential selection in chloroplast genome between *japonica* and *indica*. *Rice* 12: 65.
- Ki-Joong K, Hea-Lim L (2004) Complete Chloroplast Genome Sequences from Korean Ginseng (*Panax schinseng* Nees) and Comparative Analysis of Sequence Evolution among 17 Vascular Plants. *DNA Research* 4: 247-261.
- QW Zeng, HY Chen, C Zhang, MJ Han, T Li, et al. (2015) Definition of Eight Mulberry Species in the Genus *Morus* by Internal Transcribed Spacer-Based Phylogeny. *Plos One* 10: e0135411.
- Sancho R, Cantalapiedra CP, López-Alvarez D, Gordon SP, Vogel JP, et al. (2017) Comparative plastome genomics and phylogenomics of *Brachypodium*: flowering time signatures, introgression and recombination in recently diverged ecotypes. *New Phytol* 218: 1631-1644.
- Ghosh JS, Bhattacharya S, Pal A (2017) Molecular phylogeny of 21 tropical bamboo species reconstructed by integrating non-coding internal transcribed spacer (ITS1 and 2) sequences and their consensus secondary structure. *Genetica* 145: 319-333.
- B Xu, XM Zeng, XF Gao, DP Jin, LB Zhang (2017) ITS non-concerted evolution and rampant hybridization in the legume genus *Lespedeza* (Fabaceae). *Sci Rep* 7: 40057.
- Anton Bankevich, Sergey Nurk, Dmitry Antipov, Alexey A Gurevich, Mikhail Dvorkin, et al. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19: 455-477.
- GonçAlves NA, Monica C, Garcez LL (2015) Complete Chloroplast Genome of *Tanaecium tetragonolobum*: The First Bignoniaceae Plastome. *Plos One* 10: e0129930.
- QL Li, JZ Guo, N Yan, CC Li (2016) Complete chloroplast genome sequence of cultivated *Morus* L. species. *Genet Mol Res* 15.
- HY Zhang, C Li, HM Miao, SJ Xiong (2013b) Insights from the Complete Chloroplast Genome into the Evolution of *Sesamum indicum* L. *Plos One* 8: e80508.
- Zhengqiu Cai, Cynthia Penaflor, Jennifer V Kuehl, James Leebens-Mack, John E Carlson, et al. (2006) Complete plastid genome sequences of *Drimys*, *Liriodendron*, and *Piper*: implications for the phylogenetic relationships of magnoliids. *BMC Evol Biol* 6: 77.
- Steane DA (2005) Complete Nucleotide Sequence of the Chloroplast Genome from the Tasmanian Blue Gum, *Eucalyptus globulus* (Myrtaceae). *DNA Res* 12: 215-220.
- Katti MV, Ranjekar PK, Gupta VS (2001) Differential Distribution of Simple Sequence Repeats in Eukaryotic Genome Sequences. *Mol Biol Evol* 18: 1161-1168.
- Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am J Bot* 94: 275-288.
- Magdy M, Ou L, Yu H, Rong Chen, Yuhong Zhou, et al. (2019) Pan-plastome approach empowers the assessment of genetic variation in cultivated *Capsicum* species. *Hortic Res* 6: 108.
- Naqvi S, Godfrey AK, Hughes JF, Goodheart ML, Mitchell RN, et al. (2019) Conservation, acquisition, and functional impact of sex-biased gene expression in mammals. *Science* 365: eaaw7317.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2021.40.006418

Weiguo Zhao. Biomed J Sci & Tech Res

This work is licensed under Creative Commons Attribution 4.0 License

Submission Link: <https://biomedres.us/submit-manuscript.php>

Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

<https://biomedres.us/>