



OPEN ACCESS

EDITED BY

Miloslava Fojtova,
Masaryk University, Czechia

REVIEWED BY

Fengqing Han,
Chinese Academy of Agricultural Sciences,
China
Qiang Gao,
Xinjiang Academy of Agricultural Sciences,
China

*CORRESPONDENCE

Jianghua Huang

✉ jhhuangzk@zhku.edu.cn

Xiuchun Dai

✉ 18922129009@163.com

Hailong Ren

✉ renhailong_2006@163.com

Donglin Xu

✉ icejoy2009@icloud.com

[†]These authors have contributed equally to this work

RECEIVED 02 August 2024

ACCEPTED 25 September 2024

PUBLISHED 10 October 2024

CITATION

Xiao W, Wu X, Zhou X, Zhang J, Huang J, Dai X, Ren H and Xu D (2024) Assembly and comparative analysis of the first complete mitochondrial genome of zicaitai (*Brassica rapa* var. *Purpuraria*): insights into its genetic architecture and evolutionary relationships. *Front. Plant Sci.* 15:1475064. doi: 10.3389/fpls.2024.1475064

COPYRIGHT

© 2024 Xiao, Wu, Zhou, Zhang, Huang, Dai, Ren and Xu. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Assembly and comparative analysis of the first complete mitochondrial genome of zicaitai (*Brassica rapa* var. *Purpuraria*): insights into its genetic architecture and evolutionary relationships

Wanyu Xiao^{1†}, Xian Wu^{2†}, Xianyu Zhou¹, Jing Zhang¹, Jianghua Huang^{3*}, Xiuchun Dai^{1*}, Hailong Ren^{4*} and Donglin Xu^{1*}

¹Guangzhou Municipal Crop Seed Quality Inspection Center, Guangzhou Academy of Agricultural and Rural Sciences, Guangzhou, China, ²Northeast Agricultural University, Harbin, China, ³College of Agriculture and Biology, Zhongkai University of Agriculture and Engineering, Guangzhou, China, ⁴Guangdong Provincial Key Laboratory of Crop Genetic Improvement, Crops Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou, China

Introduction: Zicaitai (*Brassica rapa* var. *purpuraria*) is a *Brassica* variety renowned for its distinctive taste and rich nutritional profile. In recent years, the mitochondrial genomes of several *Brassica* species have been documented, but the mitogenome of Zicaitai remains unreported.

Methods: In this study, we characterized the Zicaitai mitogenome achieved through the assembly of sequencing reads derived from both the Oxford Nanopore and Illumina platforms. A detailed comparative analysis was carried out with other *Brassica* species to draw comparisons and contrasts. In-depth analyses of codon usage patterns, instances of RNA editing, and the prevalence of sequence repeats within the mitogenome were also conducted to gain a more nuanced understanding of its genetic architecture. A phylogenetic analysis was performed, utilizing the coding sequences (CDS) from the mitochondrial genome of Zicaitai and that of 20 closely related species/varieties to trace evolutionary connections.

Results: The Zicaitai mitogenome is characterized by a circular structure spanning 219,779 base pairs, and it encompasses a total of 59 genes. This gene set includes 33 protein-coding genes, 23 tRNA genes, and 3 rRNA genes, providing a rich foundation for further genomic study. An analysis of the Ka/Ks ratios for 30 protein-coding genes shared by the mitogenomes of Zicaitai and seven other *Brassica* species revealed that most of these genes had undergone purifying selection. Additionally, the study explored the migration of genes between the chloroplast and nuclear genomes and the mitogenome, offering insights into the dynamics of genetic exchange within the *Brassica* genus.

Discussion: The collective results in this study will serve as a foundational resource, aiding future evolutionary studies focused on *B. rapa*, and contributing to a broader understanding of the complexities of plant evolution.

KEYWORDS

zicaitai, *Brassica*, mitochondrial genome, phylogenetic analysis, gene transfer

1 Introduction

Brassica is a critically important genus within the *Brassicaceae* family (Essoh et al., 2020). It encompasses six principal species: *Brassica rapa* (AA, $2n = 2x = 20$), *B. oleracea* (CC, $2n = 2x = 18$), *B. nigra* (BB, $2n = 2x = 16$), *B. napus* (AACC, $2n = 4x = 38$), *B. juncea* (AABB, $2n = 4x = 36$) and *B. carinata* (BBCC, $2n = 4x = 34$) (Koh et al., 2017; Zhang et al., 2022). The genetic interrelations among these species are often depicted as the ‘Triangle of U’ (Wu et al., 2022). Zicaitai (*B. rapa* var. *purpuraria*), a variety of *B. rapa* that originated in Central China as a result of extensive cultivation and domestication (Li et al., 2019), is widely cultivated in southern China. This variety is a popular vegetable nationwide, celebrated for its distinctive flavor profile and substantial nutritional benefits, particularly because of its rich anthocyanin content in the leaves and stalks (Podsędek, 2017; Guo et al., 2015). Anthocyanins, known for their diverse biological roles in plants and their health benefits to humans (Hayashi et al., 2010; Lee et al., 2017; Zhang and Jing, 2022; Nistor et al., 2022), have propelled zicaitai into the spotlight in recent years (Guo et al., 2015; Li et al., 2019; Zhang X. et al., 2020).

Mitochondria are semi-autonomous organelles found in the majority of eukaryotic cells (Ye et al., 2017) and perform a range of essential functions, including energy conversion, tricarboxylic acid cycling, oxidative phosphorylation, calcium ion storage, cell proliferation and metabolism, etc (Hamani and Giege, 2014; Møller et al., 2021). Plant mitochondrial genomes are considerably variable in length, gene arrangement, and gene content (Birky, 1995; Bi et al., 2020), and they contain large repeats that can result in isomerization within a species. Recent studies have also highlighted the frequent gene transfers between nuclear, mitochondrial, and chloroplast genomes as a common phenomenon during plant evolution (Chang et al., 2013; Zhao et al., 2019). Given the substantial impact of plant organellar genomes on metabolism and adaptation, a detailed analysis of their sequences and structures can shed light on the evolutionary history and the intricate interplay between these genomes (Pinard et al., 2019).

Leveraging the advancements in sequencing technology, numerous *Brassica* mitochondrial genomes have been characterized, including the mitogenomes of several subspecies in *B. rapa* (Handa, 2003; Chang et al., 2011; Chen et al., 2011; Wu et al., 2019; Shao et al., 2021). However, the situation among the subspecies/varieties of *B. rapa* remains confusing, especially for the Zicaitai, whose selection and

domestication are still unknown. Studying the mitochondrial genome of Zicaitai is of significant value for understanding its evolutionary position within the species, genetic variation, phylogeny, and domestication history of the Chinese cabbage varieties. However, the mitochondrial genome of zicaitai remains largely uncharted territory. In this study, we sequenced and annotated the its mitogenome, conducting a detailed analysis of its genomic features, codon usage, RNA editing, sequence repeats, and comparative genomics. The findings provide an in-depth understanding of the zicaitai mitochondrial genome, which is expected to bolster ongoing research efforts into the evolution of *B. rapa*.

2 Materials and methods

2.1 Plant materials and DNA sequencing

Zicaitai seeds, designated ‘Youxuan hongshan’ were sourced from the Guangzhou Academy of Agricultural Sciences. Germination of the seeds was initiated at 25°C within a growth chamber, followed by transplantation into a greenhouse under a 16/8-hour light-dark photoperiod at a temperature of 25°C, for a 6-week growth. Twenty grams of zicaitai leaf tissue from one single well-grown seedling was then harvested, promptly transported using dry ice, and dispatched to Genepioneer Biotechnologies Company in Nanjing, China. The integrity of the DNA samples was assessed via agarose gel electrophoresis, and their concentrations were quantified using a Nanodrop instrument (model 2000c UV-Vis). Samples meeting the quality criteria were subsequently submitted for sequencing using both Oxford Nanopore and Illumina platforms.

2.2 Mitogenome assembly and annotation

Given the high conservation of plant mitochondrial genes, including coding sequences (CDS) and ribosomal RNA (rRNA), the minimap2 software (version 2.1) (Li, 2018), a tool designed for tri-generational comparison, was utilized. This software compared the original tri-generational data against a reference gene sequence, with preference given to sequences exceeding 50 base pairs in length as candidate sequences. The seed sequence was chosen based on its enrichment in

core genes and superior relative quality, indicating a more complete representation of the core genes. Subsequently, minimap2 was employed to align the original sequencing data with the seed sequence, and sequences with an overlap exceeding 1 kilobase and a similarity higher than 70% were integrated into the seed sequence. This iterative comparison of the original data with the seed sequence facilitated the acquisition of comprehensive mitochondrial genome data. The tri-generational assembly software Canu (Koren et al., 2017) was then applied to refine the obtained data. Bowtie2 (version 2.3.5.1) (Langmead and Salzberg, 2012) was used to align the second-generation data with the corrected sequence, followed by the application of Unicycler (version 0.4.8) (Wick et al., 2017) to align the second-generation data with the corrected tri-generational data. The final assembly was manually curated to determine the orientation of the branches. The mitochondrial genome map was generated using OGDRAW (<https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>) (Lohse et al., 2007). Mitochondrial annotation proceeded through the following steps: (1) BLAST was employed to align published plant mitochondrial sequences with their respective encoded proteins and rRNA, with subsequent manual refinements based on related species; (2) Transfer RNA (tRNA) was annotated using tRNAscanSE (<http://lowelab.ucsc.edu/tRNAscan-SE/>) (Chan and Lowe, 2019); (3) Open Reading Frames (ORFs) were annotated with the Open Reading Frame Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>), considering the shortest length of 102 base pairs. For the purpose of collinearity analysis, the nucmer (4.0.0beta2) software was executed utilizing the “-maxmatch” parameter to facilitate genomic alignments between the assembled mitochondrial sequence of zicaitai and those of related species or varieties, resulting in the creation of Dot-plot graphs.

2.3 Analysis of repeat structures and SSRs

Dispersed repeats were identified using the REPuter tool (Kurtz et al., 2001); tandem repeats were detected with the Tandem Repeats Finder software (trf409.linux64: 2 7 7 80 10 50 2000 -f -d -m) (Benson, 1999); and simple sequence repeats (SSRs) were analyzed using the MISA program (Beier et al., 2017), with a minimum distance between SSRs set at 100 base pairs and the motif size for one- to six-nucleotide SSRs set to 10, 5, 4, 3, 3, and 3, respectively.

2.4 Prediction of RNA editing sites and codon usage

The codon composition within the mitochondrial genome of zicaitai was examined using a custom Perl script, which was designed to identify unique coding sequences (CDS), determine the number of codons per gene, and calculate the Relative Synonymous Codon Usage (RSCU) for synonymous codons. RNA editing sites in the protein-coding genes (PCGs) of zicaitai and seven other Brassica mitogenomes (*Brassica carinata*, *B. juncea*, *B. napus*, *B. oleracea*, *B. rapa* subsp. *Oleifera*, *B. nigra*, and *B. rapa* subsp. *Nipposinica*) were predicted using the online resource <http://prep.unl.edu/> (Mower, 2005).

2.5 Ka/Ks analysis and phylogenetic tree construction

The nonsynonymous (Ka) and synonymous (Ks) substitution rates for each PCG between zicaitai and seven other reported Brassica species (*B. rapa* subsp. *nipposinica*, *B. rapa* subsp. *oleifera*, *B. juncea*, *B. napus*, *B. oleracea*, *B. carinata* and *B. nigra*) were calculated using DnaSP (Librado, 2009). A total of 21 complete mitogenomes (Supplementary Table S5) were utilized to determine the phylogenetic position of zicaitai. The maximum likelihood evolutionary tree was constructed using 30 conserved mitochondrial PCG genes (*atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *ccmB*, *ccmC*, *ccmFc*, *ccmFn*, *cob*, *cox1*, *cox2*, *cox3*, *matR*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *nad9*, *rpl2*, *rpl5*, *rps12*, *rps14(-)*, *rps3*, *rps4*, and *rps7*) present across the 21 analyzed species. The mafft software (version 7.427, -auto mode) was employed for multiple sequence alignment of inter-species sequences. The aligned sequences were concatenated, and trimAl (version 1.4, -0.7 gt) was used for refinement. The jmodeltest-2.1.10 software was then utilized to select the best-fit model, which was identified as the GTR model. Finally, the RAXML v8.2.10 software (<https://cme.h-its.org/exelixis/software.html>) was used with the GTRGAMMA model and a bootstrap value of 1000 to construct the maximum likelihood evolutionary tree.

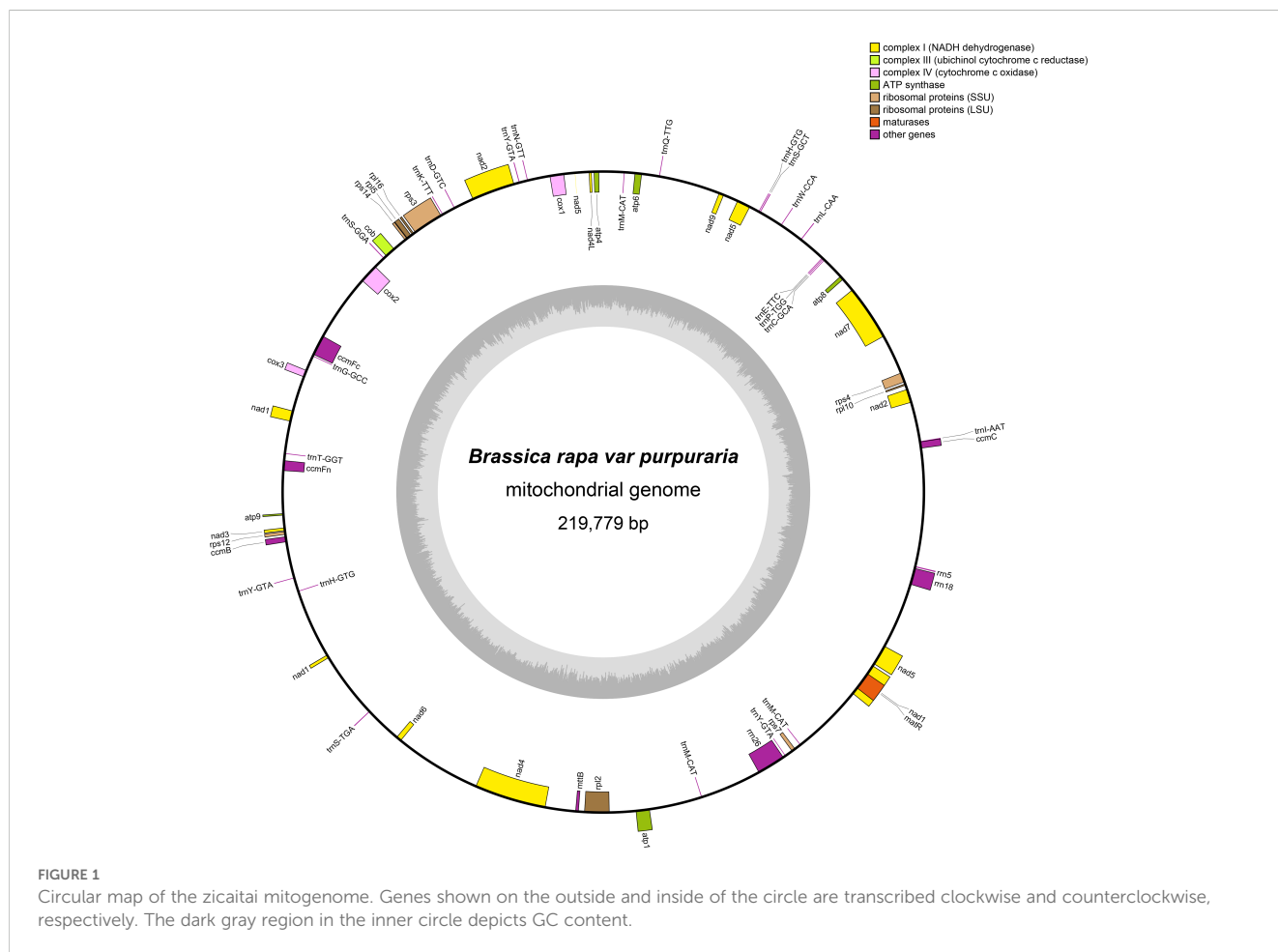
2.6 Genome alignments

The zicaitai mitogenome was compared against the chloroplast genome of zicaitai (OP729397.1) using BLASTN 2.9.0+ with the following criteria: a matching rate of $\geq 70\%$, an E-value of $\leq 1e-5$, and a sequence length of ≥ 40 base pairs (Wick et al., 2017). To identify potential regions of nuclear origin within the zicaitai mitogenome, a BLASTN search (with a maximum E-value of $1e-50$) was also conducted against all contigs from the previously sequenced zicaitai nuclear genome. Sequences longer than 250 base pairs with a pairwise similarity of $> 80\%$ were subjected to annotation in the GenBank database to identify their potential nuclear origins.

3 Results

3.1 Genomic features of the zicaitai mitogenome

The mitogenome of zicaitai, as determined through sequencing with both the Oxford Nanopore and Illumina platforms, has been deposited in the GenBank database under the accession number OP729396.1. This complete mitogenome spans a length of 219,779 base pairs (bp) and exhibits the characteristic circular conformation, as depicted in Figure 1. The mitochondrial DNA (mtDNA) nucleotide composition is comprised of adenine (A) at 27.45%, guanine (G) at 22.32%, cytosine (C) at 22.92%, and thymine (T) at 27.31%, with a GC content of 45.24% (Table 1). Protein-coding genes (PCGs), introns, tRNA genes, rRNA genes and non-coding regions account for 13.22%, 12.86%, 0.79%, 2.34% and 70.79% of the whole mitogenome, respectively.



The zicaikai mitogenome contains 977 open reading frames (ORFs) and 59 genes, including 33 protein-coding genes, 23 tRNA genes, and 3 rRNA genes (Table 2). The protein-coding genes are categorized into 9 functional groups: 5 for ATP synthases, 4 for cytochrome c biogenesis, 1 for ubiquinol cytochrome c reductase, 3 for cytochrome c oxidase, 1 for maturase, 1 for transport membrane protein, 9 for NADH dehydrogenase, 4 for LSU ribosomal proteins, and 5 for SSU ribosomal proteins. Additionally, three tRNA genes (trnH-GTG, trnM-CAT, trnY-GTA) occur in two or three copies and are located within repeat sequences (Figure 1).

Mitogenomic sequence-based collinearity analysis has uncovered a high degree of collinearity between the mitogenome of zicaikai and those of *B. rapa subsp. nipposinica*, *B. rapa subsp. oleifera*, and notably, *B. juncea*, which is a distinct species from zicaikai, exhibiting minimal genomic rearrangement characteristics. In contrast, substantial rearrangements were observed when comparing zicaikai with other Brassica species (Supplementary Figure S1).

3.2 Codon usage analysis of PCGs

The zicaikai mitogenome's protein-coding genes (PCGs) span a total of 29,055 bp. The majority of these PCGs initiate with the

standard ATG start codon and terminate with the common stop codons TAA, TGA, or TAG. Exceptions include the *nad1* gene with an ACG start codon and the *ccmFc* gene with a CGA stop codon; both exhibit the phenomenon of C to U RNA editing (Table 2).

Additionally, we examined the relative synonymous codon usage (RSCU) for the 33 PCGs in the zicaikai mitogenome (Figure 2). Excluding termination codons, these 33 PCGs encode for 9,652 codons. The analysis indicated that leucine (Leu) (10.91%), serine (Ser) (8.87%), and arginine (Arg) (6.48%) are the most frequently used amino acids, as shown in Supplementary Table S1. The RSCU values for nearly all NNT and NNA codons exceeded 1.0, with the exceptions of Ala (GCA, 0.96), Thr (ACA, 0.91), Leu (CUA, 0.89), and Ile (AUA, 0.84). A strong bias toward T (U) or A at the third position of the codons in the zicaikai can be observed, a pattern commonly found in land plant mitogenomes (Ma et al., 2022).

3.3 Analysis of synonymous and nonsynonymous substitution ratios

In the field of genetics, the ratio of nonsynonymous to synonymous substitutions (K_a/K_s) serves as a valuable tool for

TABLE 1 Nucleobase constitution of the zicaitai mitogenome.

Category of sequence	Base composition (%)					Size in bp (proportion in percentage)
	A	G	C	T	GC	
Whole genome	27.45	22.32	22.92	27.31	45.24	219,779 (100%)
Protein-coding genes	26.34	21.7	20.92	31.03	42.63	29,055 (13.22%)
Introns	25.61	26.86	24.22	23.31	51.08	28,253 (12.86%)
tRNA genes	22.57	28.76	22.86	25.81	51.62	1,728 (0.79%)
rRNA genes	26.63	28.79	22.59	21.99	51.38	5,144 (2.34%)
Non-coding regions	27.9	21.97	22.51	27.62	44.48	155,599 (70.79%)

TABLE 2 Gene profile and organization of the zicaitai mitogenome.

Gene category	Gene name	Length	Start codon	Stop codon	Amino acid
ATP synthase	<i>atp1</i>	1524	ATG	TAG	508
	<i>atp4</i>	579	ATG	TAA	193
	<i>atp6</i>	786	ATG	TAA	262
	<i>atp8</i>	477	ATG	TGA	159
	<i>atp9</i>	225	ATG	TGA	75
Cytochrome c biogenesis	<i>ccmB</i>	621	ATG	TAA	207
	<i>ccmC</i>	744	ATG	TGA	248
	<i>ccmFc</i>	1329	ATG	CGA(TGA)	443
	<i>ccmFn</i>	1146	ATG	TGA	382
Ubiquinol cytochrome c reductase	<i>cob</i>	1182	ATG	TGA	394
Cytochrome c oxidase	<i>cox1</i>	1584	ATG	TAA	528
	<i>cox2</i>	783	ATG	TAA	261
	<i>cox3</i>	798	ATG	TGA	266
Maturases	<i>matR</i>	1974	ATG	TAG	658
Transport membrane protein	<i>mttB</i>	360	ATG	TAG	120
NADH dehydrogenase	<i>nad1</i>	978	ACG(ATG)	TAA	326
	<i>nad2</i>	1467	ATG	TAA	489
	<i>nad3</i>	357	ATG	TAA	119
	<i>nad4</i>	1488	ATG	TGA	496
	<i>nad4L</i>	303	ATG	TAA	101
	<i>nad5</i>	2010	ATG	TAA	670
	<i>nad6</i>	618	ATG	TAA	206
	<i>nad7</i>	1185	ATG	TAG	395
	<i>nad9</i>	573	ATG	TAA	191
Ribosomal proteins (LSU)	<i>rpl10</i>	225	ATG	TAG	75
	<i>rpl16</i>	249	ATG	TAA	83
	<i>rpl2</i>	1050	ATG	TGA	350
	<i>rpl5</i>	558	ATG	TAA	186

(Continued)

TABLE 2 Continued

Gene category	Gene name	Length	Start codon	Stop codon	Amino acid
Ribosomal proteins (SSU)	<i>rps12</i>	378	ATG	TGA	126
	<i>rps14</i>	303	ATG	TAG	101
	<i>rps3</i>	1665	ATG	TAG	555
	<i>rps4</i>	1089	ATG	TAA	363
	<i>rps7</i>	447	ATG	TAA	149
Ribosomal RNAs	<i>rrn18</i>	1847	-	-	-
	<i>rrn26</i>	3176	-	-	-
	<i>rrn5</i>	121	-	-	-
Transfer RNAs	<i>trnC-GCA</i>	71	-	-	-
	<i>trnD-GTC</i>	74	-	-	-
	<i>trnE-TTC</i>	72	-	-	-
	<i>trnG-GCC</i>	72	-	-	-
	<i>trnH-GTG (2)</i>	74/74	-	-	-
	<i>trnI-AAT</i>	69	-	-	-
	<i>trnK-TTT</i>	73	-	-	-
	<i>trnL-CAA</i>	85	-	-	-
	<i>trnM-CAT (3)</i>	73/74/74	-	-	-
	<i>trnN-GTT</i>	72	-	-	-
	<i>trnP-TGG</i>	75	-	-	-
	<i>trnQ-TTG</i>	72	-	-	-
	<i>trnS-GCT</i>	88	-	-	-
	<i>trnS-GGA</i>	87	-	-	-
	<i>trnS-TGA</i>	87	-	-	-
	<i>trnT-GGT</i>	69	-	-	-
	<i>trnW-CCA</i>	74	-	-	-
<i>trnY-GTA (3)</i>	83/68/68	-	-	-	

Numbers after gene names are the number of copies.

elucidating the evolutionary dynamics of genes. In this study, we calculated the Ka/Ks ratio for 30 protein-coding genes (PCGs) shared by zicaitai and seven other Brassica species (*B. rapa subsp. nipposinica*, *B. rapa subsp. oleifera*, *B. juncea*, *B. napus*, *B. oleracea*, *B. carinata* and *B. nigra*) (Figure 3; Supplementary Table S2). The PCGs were found to be highly homologous among zicaitai, *B. rapa subsp. nipposinica*, *B. rapa subsp. oleifera*, and *B. oleracea*, as evidenced by a Ka/Ks ratio of zero for these 30 PCGs. Moreover, the majority of Ka/Ks ratios were below 1.0, implying that these genes have been predominantly under the influence of purifying selection throughout their evolutionary history. In contrast, a single gene, *rpl2*, exhibited a Ka/Ks ratio exceeding 1.0, indicating that it has been subject to positive selection. Lastly, two genes, *atp6* and *nad2*, displayed Ka/Ks ratios approaching 1, suggesting that they have undergone neutral evolution since the time of their common ancestor's divergence.

3.4 Analysis of repeats in the zicaitai

An analysis of long repeats in the zicaitai mitogenome identified 252 dispersed repeats, comprising 108 forward (42.86%) and 144 palindromic (57.14%) repeats; no reverse or complementary repeats were detected. The total length of these long repeats was 16,251 bp, accounting for 7.39% of the length of the complete mitogenome. The average length of the repeats was 64.49 bp, with the longest repeat spanning 2,427 bp. The majority of the repeats (153 repeats, 60.71%) were within the range of 30 to 50 bp in length (Supplementary Table S3).

Tandem repeat sequences, also known as satellite DNAs, are formed by the sequential arrangement of short sequences that serve as repeating units, typically ranging from 1 to 200 bases in length. As detailed in Table 3, a total of 12 tandem repeats were identified in the zicaitai mitogenome, with lengths varying from 12 to 39 bp, and all of them exhibited a match degree exceeding 95%.

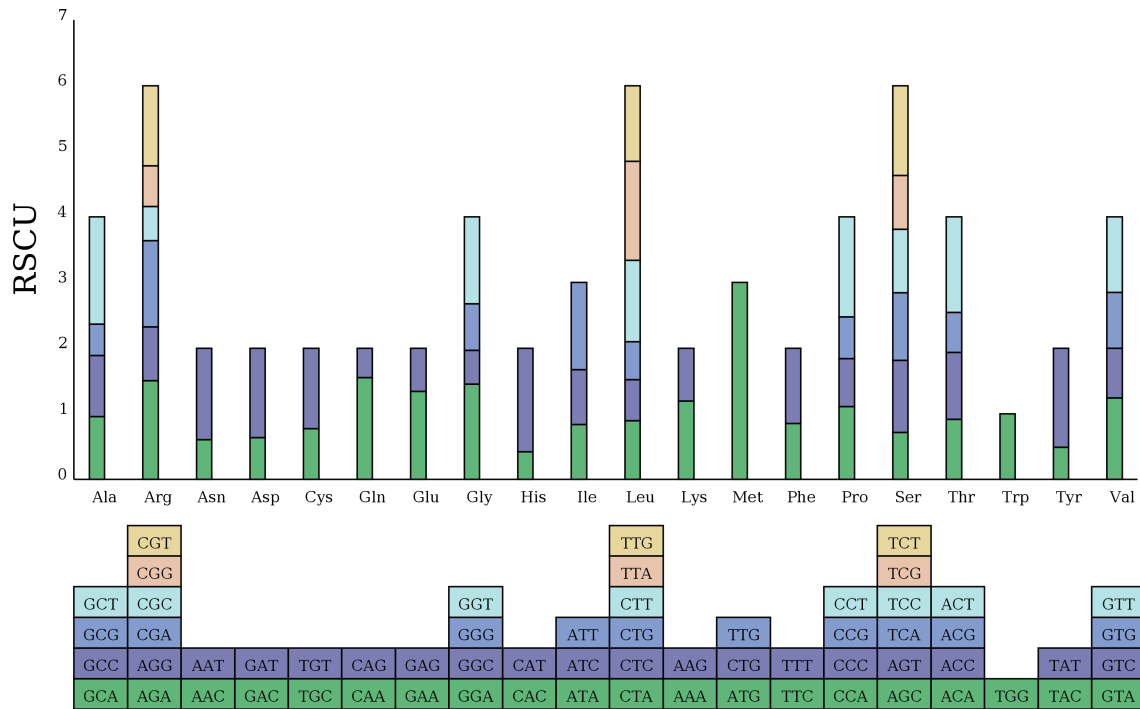


FIGURE 2 Relative synonymous codon usage (RSCU) in the zicaitai mitogenome. Codon families are shown on the x-axis. RSCU values represent the observed frequency of a specific codon in relation to the expected frequency if synonymous codons were used uniformly.

Microsatellites, also known as simple sequence repeats (SSRs), are DNA fragments ranging from 1 to 6 bp in length and are useful molecular markers for studying genetic diversity and identifying species (Ma et al., 2022). In our research, we identified 55 SSRs in the zicaitai mitogenome, comprising 20 mono-, 11 di-, 5 tri-, 18 tetra-, and 1 pentanucleotide repeat (Table 4). Over 89% of these

SSRs were mono-, di-, or pentanucleotide types. Detailed analysis showed that the A/T repeat accounted for 32.73% (18/55) of the SSRs, followed by the AG/CT repeat at 12.73% (7/55) and the AATG/ATTC repeat at 9.09% (5/55). This is different from the most abundant SSR types in *B. rapa*'s nuclear genome, where A/T and AT/TA repeats are predominant (Gao et al., 2022).

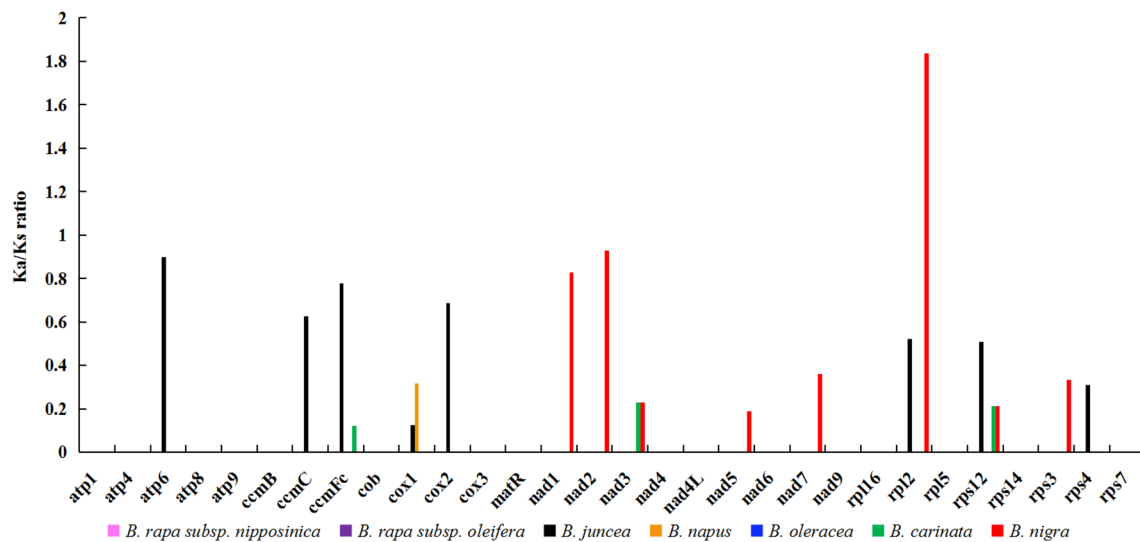


FIGURE 3 Ka/Ks ratios of 30 protein-coding genes in zicaitai in comparison with *B. rapa subsp. nipposinica*, *B. rapa subsp. Oleifera*, *B. juncea*, *B. napus*, *B. oleracea*, *B. carinata* and *B. nigra*, respectively.

TABLE 3 Tandem repeat sequences in zicaitai mitogenome.

Repeat sequence	Size (bp)	Start site	End site	Percent Matches	Number of copy
AAAGGAGAGGTGCTTTAGCAACTCGACTG	29	28,136	28,225	98	3.1
GCTTCTTGTTGTTGATGAGCTTATACAC	27	49,103	49,176	95	2.7
TGAAGTATAGC	12	61,926	61,950	100	2.1
TCGAGATCTTTGAACCTTTCAG	22	69,116	69,164	96	2.2
ATGTTAGTGTTCAGTATATC	20	71,738	71,775	100	1.9
CTCGAGGAACGC	12	88,938	88,964	100	2.2
CGGAGGCGGTAAG	14	116,463	116,492	100	2.1
AGATTTTACAAATGGTCTG	19	118,981	119,019	100	2.1
TATCAATTTTATAAGAGAAGAAAGATCGTTTTTTTAAAT	39	119,161	119,271	100	2.8
ATATATCCATTCTCATA	17	138,234	138,272	95	2.3
TAGAAAACCTGGCATC	15	147,084	147,113	100	2.0
AACAGATAACAACAGCATATT	21	205,177	205,233	100	2.7

3.5 Prediction of RNA editing sites in PCGs

A total of 378 RNA editing sites were identified in the zicaitai mitogenome, all featuring C-to-U changes (Table 5). Of these sites, 30.42% (115/378) were located at the first position and 65.34% (247/378) at the second position of codons, while a small fraction, 4.23% (16/378), were found at both positions. There was significant variation in the number of RNA editing sites across different genes, with the highest counts observed in the NADH dehydrogenase genes (*nad4*, *nad5*, *nad1*, *nad2*, and *nad7*) and those involved in cytochrome c biogenesis (*ccmB* and *ccmC*). In contrast, no RNA editing sites were detected in the *atp8* and *cox1* genes of zicaitai (Supplementary Table S4).

We examined RNA editing sites across 33 protein-coding genes (PCGs) present in the mitogenomes of zicaitai and seven other *Brassica* species (Figure 4), and predicted a total of 2,823 RNA editing sites, with counts varying from 308 in *B. nigra* to 397 in both *B. rapa subsp. Oleifera* and *B. oleracea*. Zicaitai and *B. juncea* had the most similar counts, with 378 and 375 sites respectively.

3.6 Phylogenetic analysis

To ascertain the phylogenetic position of zicaitai, we retrieved 20 plant mitogenomes from GenBank (<https://www.ncbi.nlm.nih.gov/genome/browse/>) and constructed a phylogenetic tree using a set of 30 conserved single-copy orthologous genes that are present across all 21 analyzed mitogenomes (Supplementary Table S5). As depicted in Figure 5, 13 out of 18 nodes within the constructed tree had bootstrap support values exceeding 70%, with 9 nodes achieving full support at 100%. The phylogenetic analysis robustly endorses the close phylogenetic ties between species in genus *Brassica*. Furthermore, it also indicated a close relationship between *B. rapa* and *B. napus*, aligning with the findings of Chang et al. (2011), who suggested that *B. napus* possesses a primary mitotype highly similar to that of *B. rapa*. Collectively, our mitogenomic analysis lays a valuable groundwork for future studies on the phylogenetic relationships within the *Brassica* genus.

TABLE 4 Frequencies of identified SSR motifs in zicaitai mitogenome.

Motif type	Number of repeats										Total	Proportion (%)
	3	4	5	6	8	10	11	13	14	16		
Monomer						14	2	1	1	2	20	36.4
Dimer			10	1							11	20.0
Trimer		4			1						5	9.1
Tetramer	17	1									18	32.7
Pentamer	1										1	1.8
Total	18	5	10	1	1	14	2	1	1	2	55	100.0

TABLE 5 Prediction of RNA editing sites.

Type	RNA editing	Number	Percentage
hydrophilic	CAC (H) → TAC (Y)	6	12.70%
	CAT (H) → TAT (Y)	16	
	CGC (R) → TGC (C)	6	
	CGT (R) → TGT (C)	20	
hydrophilic-hydrophobic	ACA (T) → ATA (I)	2	45.77%
	ACC (T) → ATC (I)	1	
	ACG (T) → ATG (M)	6	
	ACT (T) → ATT (I)	4	
	CGG (R) → TGG (W)	21	
	TCA (S) → TTA (L)	53	
	TCC (S) → TTC (F)	16	
	TCG (S) → TTG (L)	38	
	TCT (S) → TTT (F)	32	
hydrophilic-stop	CGA (R) → TGA (X)	1	0.26%
hydrophobic-hydrophilic	CCA (P) → TCA (S)	6	8.73%
	CCC (P) → TCC (S)	5	
	CCG (P) → TCG (S)	5	
	CCT (P) → TCT (S)	17	
hydrophobi	CCA (P) → CTA (L)	34	32.54%
	CCC (P) → CTC (L)	7	
	CCC (P) → TTC (F)	6	
	CCG (P) → CTG (L)	20	
	CCT (P) → CTT (L)	23	
	CCT (P) → TTT (F)	10	
		4	

(Continued)

TABLE 5 Continued

Type	RNA editing	Number	Percentage
	CTC (L) → TTC (F)	8	
	CTT (L) → TTT (F)		
	GCA (A) → GTA (V)		
	GCC (A) → GTC (V)		
	GCG (A) → GTG (V)		
	GCT (A) → GTT (V)		
Total	—	378	100%

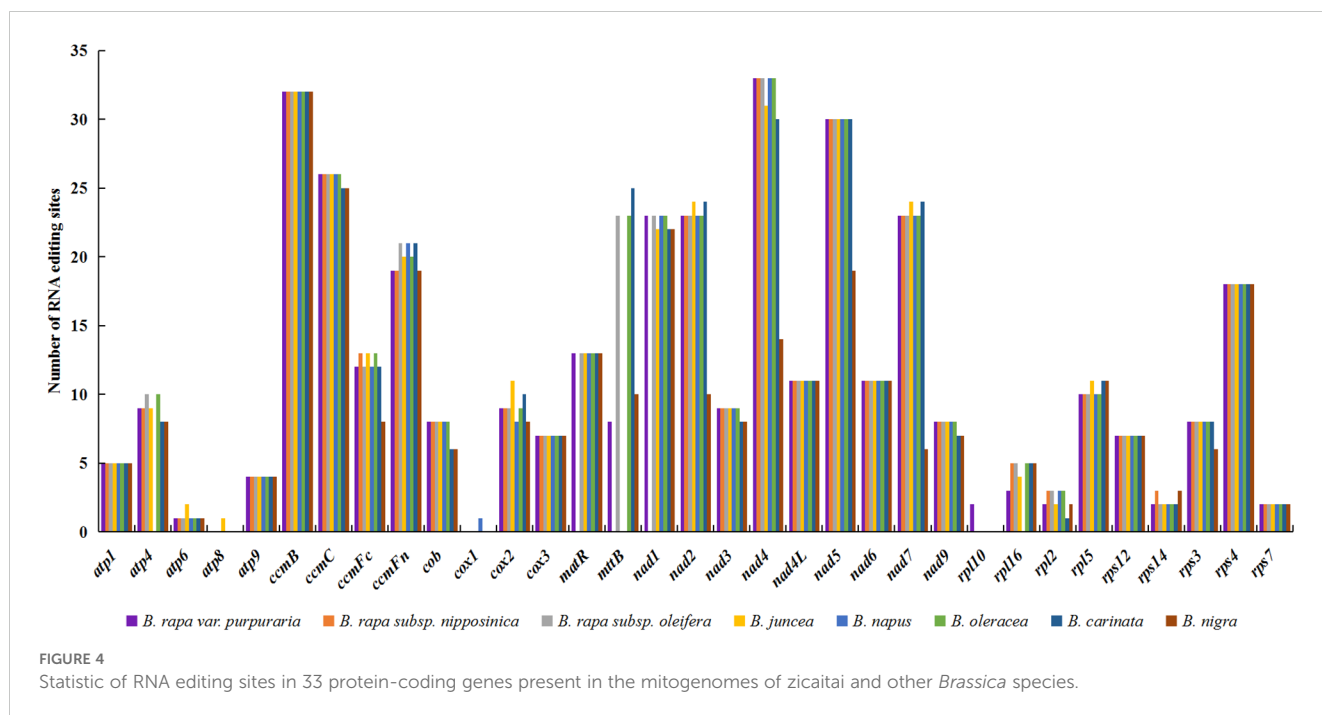
3.7 Sequence transfer events between the nuclear and organellar genomes in zicaitai

There are six potential directions for gene transfer between nuclear, mitochondrial and chloroplast genomes during plant evolution. To gain a deeper understanding of the sequence transfer events in zicaitai, we conducted a search using the zicaitai mitogenome sequences as queries against the zicaitai nuclear and chloroplast genomes. Our search yielded 1040 hits, encompassing 267,131 base pairs (bp) of sequences shared by both the nuclear genome and the mitogenome (Supplementary Table S6). Based on the alignment between the nuclear and mitochondrial genomes, we observed hits across every chromosome of zicaitai (Figure 6). However, the total lengths of these hits and their respective coverage percentages varied among the chromosomes. Notably, Chromosome 6 exhibited the longest cumulative length of hits (137,674 bp), significantly surpassing the lengths found on the other chromosomes.

The zicaitai mitogenome sequence, spanning 219,779 bp, is approximately 1.43 times the length of its chloroplast genome (153,621 bp). We identified 22 fragments totaling 13,323 bp, representing 6.06% of the mitogenome, that appear to be the result of gene exchange between the chloroplast genome and the mitogenome of zicaitai (Table 6). These fragments encompass seven complete chloroplast genes: *ycf15*, *trnL-CAA*, *trnN-GUU*, *trnW-CCA*, *trnD-GUC*, *trnM-CAU*, and *trnI-CAU*. The remaining fragments consist of partial gene sequences or intergenic spacer regions in the chloroplast genome. Interestingly, our analysis revealed that DNA migration frequently occurred within the inverted repeat region of the zicaitai chloroplast genome.

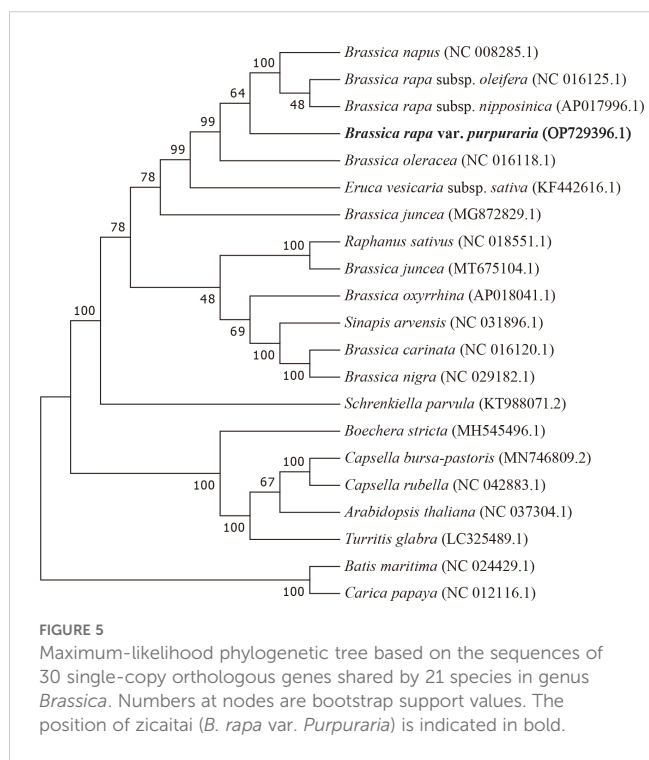
4 Discussion

Mitochondria are the energy-producing factories essential for life processes. Plant mitogenomes have more complex structures than those of animals due to size variation, the presence of repeated



sequences and other factors (Kozik et al., 2019; Cheng et al., 2021). The availability of rapid and cost-effective genome sequencing technologies has significantly advanced our understanding of mitogenomes. In this study, we characterized for the first time the complete mitogenome of zicaitai, a distinctive *B. rapa* variety. We found that the zicaitai mitogenome has a similar size and GC content to that of other *Brassica* species (Handa, 2003; Chang et al., 2011) and exhibits high degree of collinearity with the mitogenomes

of *B. rapa subsp. nipposinica*, *B. rapa subsp. oleifera*, and a distinct species, *B. juncea*, distinguishing zicaitai from other *Brassica* species (Supplementary Figure S1). Similar to most other mitogenomes, most sequences in the zicaitai mitogenome are non-coding. The zicaitai mitogenome, like many others, predominantly consists of non-coding sequences, with protein-coding genes (PCGs) making up 13.22%. This is similar to *B. rapa subsp. nipposinica* (12.36%), *B. juncea* (14.85%), and *B. oleracea* (13.44%), yet significantly lower than the 20-23% found in other closely related *Brassica* species/varieties. Intriguingly, zicaitai, although belonging to the same species as *B. rapa subsp. oleifera*, exhibits a lower proportion of protein-encoding sequences compared to the latter (22.99%), suggesting an ongoing process of genetic differentiation within the species, likely driven by domestication.



Our research showed that in the zicaitai mitogenome, all protein-coding genes (PCGs) initiate with the standard ATG start codon, with the sole exception of the *nad1* gene, which starts with ACG - a common occurrence in angiosperms due to RNA editing (Ye et al., 2017; Bi et al., 2020; Ma et al., 2022). In plants, RNA editing is crucial for gene expression, particularly with the cytidine (C) to uridine (U) transition, which is prevalent in mitochondrial and chloroplast genomes (Ma et al., 2022). We also identified the *cox1* gene, which is known to involve in horizontal gene transfer among angiosperms and vary in its copy number across species as well as populations within a species (Zhang C. et al., 2020). Further studies are needed to clarify the role played by the *cox1* gene in *Brassica* evolution. A single gene (*rpl2*) was discovered in this study to present a Ka/Ks ratio exceeding 1.0, consistent with the previous research on other mitochondrial genes with similarly elevated ratios (Ye et al., 2017; Bi et al., 2016, 2020). The significance of the *rpl2* gene in gene selection and the evolutionary process of *Brassica* species/varieties warrants further investigation.

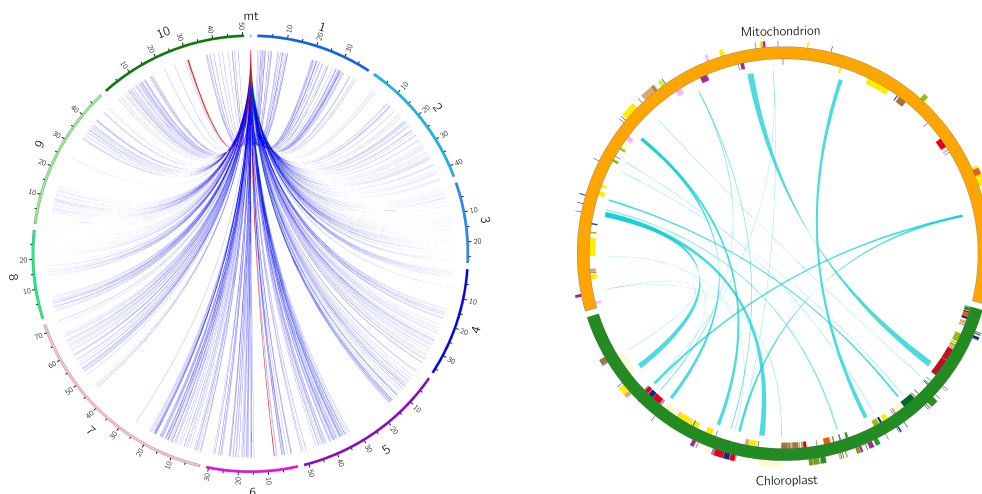


FIGURE 6 Homologous fragments identified between zicaitai mitogenome and chromosomes (left) and between the mitogenome and chloroplast genome (right). Zicaitai’s ten chromosomes are labeled numerically from 1 to 10 on the left circle. Both circles depict the linkage of homologous fragments across various genome types, connected by curves.

TABLE 6 Fragments found to be shared by zicaitai’s chloroplast genome and its mitogenome.

No.	Alignment Length	Identity (%)	Mismatch (bp)	Gap opens	CP Start	CP End	Mt Start	Mt End	gene
1	2186	98.81	26	0	23007	25192	108052	110237	<i>rpoB</i>
2	1878	97.60	30	5	91634	93500	33491	31618	<i>ycf2, ycf15, trnL-CAA</i>
3	1878	97.60	30	5	143266	145132	31618	33491	<i>trnL-CAA, ycf15, ycf2</i>
4	1365	96.41	39	3	53575	54934	142593	141234	<i>rbcL</i>
5	1159	99.66	4	0	108069	109227	63318	62160	<i>trnN-GUU, ycf1</i>
6	1159	99.66	4	0	127539	128697	62160	63318	<i>ycf1, trnN-GUU</i>
7	659	96.06	20	3	40170	40826	38183	37529	<i>psaA</i>
8	229	96.94	6	1	103753	103980	89674	89446	<i>rrn23</i>
9	229	96.94	6	1	132786	133013	89446	89674	<i>rrn23</i>
10	878	74.37	170	43	135617	136469	210654	209807	<i>rrn16</i>
11	878	74.37	170	43	100297	101149	209807	210654	<i>rrn16</i>
12	91	98.90	1	0	38693	38783	48931	49021	<i>psaB, psaA</i>
13	143	86.71	11	8	65123	65264	34480	34345	<i>trnW-CCA</i>
14	74	97.30	2	0	29508	29581	71852	71779	<i>trnD-GUC</i>
15	118	85.59	17	0	65457	65574	34129	34012	<i>trnP-UGG</i>
16	77	92.21	6	0	50951	51027	52690	52614	<i>trnM-CAU</i>
17	43	100	0	0	100387	100429	117451	117409	<i>rrn16</i>
18	43	100	0	0	136337	136379	117409	117451	<i>rrn16</i>
19	76	85.53	7	3	85434	85508	5573	5501	<i>trnI-CAU</i>
20	76	85.53	7	3	151258	151332	5501	5573	<i>trnI-CAU</i>
21	42	97.62	1	0	134348	134389	17596	17555	<i>trnI-GAU</i>
22	42	97.62	1	0	102377	102418	17555	17596	<i>trnI-GAU</i>

Repeats, encompassing tandem, short, and large types, are crucial for identifying markers in population and evolutionary studies and are prevalent in mitogenomes (Bi et al., 2016; Ma et al., 2017; Liu et al., 2021). They play key roles in intermolecular recombination within mitochondrial DNA, leading to structural variations and diverse mitogenome sizes (Alverson et al., 2010; Guo et al., 2016). In this study, we found an abundance of repeats in the zicaitai mitogenome, notably one large repeat of 2,427 bp. This repeat is identical in size to the one found in the *B. oleracea* mitogenome, suggesting a close genetic link between *B. rapa* and *B. oleracea*, as depicted in Figure 5. Our discovery of SSRs in zicaitai's organelle genome provides a valuable complement to Gao et al. (2022)'s research, which identified 173,892 SSRs within the nuclear genome of *B. rapa*.

Despite variations in RNA editing site counts across zicaitai and various *Brassica* species or varieties, genes for NADH dehydrogenase and cytochrome c biogenesis consistently exhibit the highest editing frequencies among these species/varieties (Figure 6). This pattern mirrors observations from previous studies on the mitogenomes of other plant species (Ma et al., 2022). Furthermore, all identified RNA editing sites in the zicaitai mitogenomes are situated at the first and/or second codon positions. It has been hypothesized by previous researchers that the absence of editing sites at the third codon position may be attributed to the limitations of the PREP-Mt predictive method, rather than their non-existence (Mower, 2005; Bi et al., 2020). Consequently, further investigation employing experimental techniques is warranted to clarify this issue.

Previous research has identified the predominant gene transfer in angiosperms as from organellar to nuclear genomes, with secondary importance given to transfers between nuclear, plastid, and mitochondrial genomes (Martin et al., 1998; Rice et al., 2013; Zhao et al., 2018, 2019; Grewe et al., 2014). In higher plants, the total length of transferred DNA varies from 50 kb in *Arabidopsis thaliana* to 1.1 Mb in rice (Smith et al., 2011). Our study has identified a cumulative 267 kb of DNA sequences transferred from 10 chromosomes into the zicaitai mitogenome, with variations in total length and coverage among the chromosomes. Chromosome 6 contributes the largest portion, amounting to 138 kb, which is 62.6% of the mitogenome's total length. In contrast, sequences from the chloroplast genome make up only 6.1% of the zicaitai mitogenome. Future studies should explore the coevolution of this chromosome with the mitogenome and its potential implications.

In this study, we found 11 tRNA and 3 rRNA genes common to both the nuclear and mitochondrial genomes in zicaitai (Supplementary Table S6) and six tRNA genes shared by the chloroplast genome and the mitogenome. While nuclear-to-mitogenome transfers of tRNA genes have been documented in other plants (Chang et al., 2013; Ma et al., 2022), and chloroplast-to-mitogenome transfers are typical in angiosperms (Bergthorsson et al., 2003; Cheng et al., 2021; Bi et al., 2016), the more extensive migration of these genes in zicaitai suggests a more fluid genetic exchange between different cellular compartments than previously

thought. Future research should delve into the mechanisms and evolutionary implications of this genetic fluidity.

5 Conclusions

We for the first time conducted a comprehensive characterization of the zicaitai mitogenome, which spans a length of 219,779 bp and encompasses a total of 59 genes, including 33 protein-coding genes, 23 tRNA genes, and 3rRNA genes. The protein-coding genes (PCGs), cis introns, tRNA genes, and rRNA genes constitute 13.22%, 12.86%, 0.79% and 2.34%, of the zicaitai mitogenome, respectively. The Ka/Ks analysis showed that among 30 shared mitochondrial PCGs in zicaitai and other *Brassica* species, most have experienced purifying selection. Only the *rpl2* gene showed signs of positive selection, and two genes, *atp6* and *nad2*, exhibited neutral evolution. Phylogenetic analysis based on these 30 PCGs from zicaitai and 20 related species/varieties indicated that *B. rapa* is most closely related to *B. napus*, with *B. oleracea* being the next closest. Furthermore, we detected the migration of genes between the chloroplast and nuclear genomes and the zicaitai mitogenome. The collective findings in this study offer an in-depth understanding of the zicaitai mitogenome, which is instrumental for advancing evolutionary studies within the species *B. rapa*.

Data availability statement

The original mitochondrial genome presented in the study are publicly available. This data can be found in NCBI (<https://www.ncbi.nlm.nih.gov/>) under the GenBank: OP729396.1 (<https://www.ncbi.nlm.nih.gov/nucleotide/OP729396.1>). The data are publicly available. The datasets presented in this study can be found in NCBI.

Author contributions

WX: Data curation, Formal analysis, Investigation, Writing – original draft. XW: Data curation, Investigation, Writing – original draft. XZ: Investigation, Writing – original draft. JZ: Formal analysis, Funding acquisition, Writing – original draft. JH: Data curation, Writing – review & editing. XD: Funding acquisition, Supervision, Writing – original draft. HR: Conceptualization, Data curation, Writing – original draft. DX: Conceptualization, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work is supported by Guangdong Seed Industry Revitalization Project

(2022-NJS-03-001), Guangdong Vegetable Germplasm Resource Bank (Nursery) Operation and Maintenance (2023-KYXM221), and Quality Supervision and Inspection of Crop Seeds in Guangzhou City in 2024 (2024-KYXM726).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2024.1475064/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Collinearity analysis of mitogenomic sequences between zicaitai and some related species/varieties. In each box, the horizontal axis represents the assembled zicaitai's mitogenomic sequence, and the vertical axis represents the that of related species/varieties. The red lines within the boxes indicate forward alignments, while the blue lines indicate reverse complement alignments.

References

- Alverson, A. J., Wei, X., Rice, D. W., Stern, D. B., Barry, K., and Palmer, J. D. (2010). Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol. Biol. Evol.* 27, 1436–1448. doi: 10.1093/molbev/msq029
- Beier, S., Thiel, T., Münch, T., Scholz, U., and Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33, 2583–2585. doi: 10.1093/bioinformatics/btx198
- Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27, 573–580. doi: 10.1093/nar/27.2.573
- Berghthorsson, U., Adams, K., Thomason, B., and Palmer, J. (2003). Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature* 424, 197–201. doi: 10.1038/nature01743
- Bi, C., Lu, N., Xu, Y., and He, C. (2020). Characterization and analysis of the mitochondrial genome of common bean (*Phaseolus vulgaris*) by comparative genomic approaches. *Int. J. Mol. Sci.* 21, 3778. doi: 10.3390/ijms21113778
- Bi, C., Paterson, A., Wang, X., Xu, Y., Wu, D., Qu, Y., et al. (2016). Analysis of the complete mitochondrial genome sequence of the diploid cotton *Gossypium raimondii* by comparative genomics approaches. *BioMed. Res. Int.* 2016, 5040598. doi: 10.1155/2016/5040598
- Birky, C. (1995). Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *PNAS* 92, 11331–11338. doi: 10.1073/pnas.92.25.11331
- Chan, P. P., and Lowe, T. M. (2019). tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* 1962, 1–14. doi: 10.1007/978-1-4939-9173-0_1
- Chang, S., Wang, Y., Lu, J., Gai, J., Li, J., Chu, P., et al. (2013). The mitochondrial genome of soybean reveals complex genome structures and gene evolution at intercellular and phylogenetic levels. *PLoS One* 8, e56502. doi: 10.1371/journal.pone.0056502
- Chang, S., Yang, T., Du, T., Huang, Y., Chen, J., Yan, J., et al. (2011). Mitochondrial genome sequencing helps show the evolutionary mechanism of mitochondrial genome formation in *Brassica*. *BMC Genomics* 12, 497. doi: 10.1186/1471-2164-12-497
- Chen, J., Guan, R., Chang, S., Du, T., Zhang, H., and Xing, H. (2011). Substoichiometrically different mitotypes coexist in mitochondrial genomes of *Brassica napus* L. *PLoS One* 6, e17662. doi: 10.1371/journal.pone.0017662
- Cheng, Y., He, X., Priyadarshani, S., Wang, Y., Ye, L., and Qin, Y. (2021). Assembly and comparative analysis of the complete mitochondrial genome of *Suaeda glauca*. *BMC Genomics* 22, 167. doi: 10.1186/s12864-021-07490-9
- Essoh, A. P., Monteiro, F., Pena, A. R., Pais, M. S., Moura, M., and Romeiras, M. M. (2020). Exploring glucosinolates diversity in Brassicaceae: a genomic and chemical assessment for deciphering abiotic stress tolerance. *Plant Physiol. Biochem.* 150, 151–161. doi: 10.1016/j.plaphy.2020.02.032
- Gao, Y., Li, N., Li, X., Lu, Y., Feng, D., Zhang, X., et al. (2022). Genome-wide development and utilization of Simple Sequence Repeats in Chinese cabbage (*Brassica rapa* L.ssp. *pekinensis*). *Vegetable Res.* 2, 9. doi: 10.48130/VR-2022-0009
- Grewe, F., Edger, P. P., Keren, I., Sultan, L., Pires, J. C., Ostersetzer-Biran, O., et al. (2014). Comparative analysis of 11 Brassicales mitochondrial genomes and the mitochondrial transcriptome of *Brassica oleracea*. *Mitochondrion* 19 Part B, 135–143. doi: 10.1016/j.mito.2014.05.008
- Guo, N., Wu, J., and Zheng, S. (2015). Anthocyanin profile characterization and quantitative trait locus mapping in zicaitai (*Brassica rapa* L. ssp. *chinensis* var. *purpurea*). *Mol. Breed.* 35, 113. doi: 10.1007/s11032-015-0237-1
- Guo, W., Felix, G., Fan, W., and Young, G. (2016). Ginkgo and Welwitschia Mitogenomes reveal extreme contrasts in gymnosperm mitochondrial evolution. *Mol. Biol. Evol.* 33, 1448–1460. doi: 10.1093/molbev/msw024
- Hamani, K., and Giege, P. (2014). RNA metabolism in plant mitochondria. *Trends Plant Sci.* 19, 380–389. doi: 10.1016/j.tplants.2013.12.008
- Handa, H. (2003). The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Res.* 31, 5907–5916. doi: 10.1093/nar/gkg795
- Hayashi, K., Matsumoto, S., Tsukazaki, H., and Kondo, T. (2010). Mapping of a novel locus regulating anthocyanin pigmentation in *Brassica rapa*. *Breed. Sci.* 60, 76–80. doi: 10.1270/jsbbs.60.76
- Koh, J. C. O., Barbulescu, D. M., Norton, S., Redden, B., Salisbury, P. A., Kaur, S., et al. (2017). A multiplex PCR for rapid identification of *Brassica* species in the triangle of U. *Plant Methods* 13, 49. doi: 10.1186/s13007-017-0200-8
- Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116
- Kozik, A., Rowan, B., Lavelle, D., Berke, L., Schranz, M., and Michelmore, R. (2019). The alternative reality of plant mitochondrial DNA: one ring does not rule them all. *PLoS Genet.* 15, e1008373. doi: 10.1371/journal.pgen.1008373
- Kurtz, S., Choudhuri, J., Enno, O., Chris, S., Jens, S., and Robert, G. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29, 4633–4642. doi: 10.1093/nar/29.22.4633
- Langmead, B., and Salzberg, S. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Lee, Y. M., Yoon, Y., Yoon, H., Park, H. M., Song, S., and Yeum, K. J. (2017). Dietary anthocyanins against obesity and inflammation. *Nutrients* 9, 1089. doi: 10.3390/nu9101089
- Li, G. H., Chen, H. C., Liu, J. L., Luo, W. L., Xie, D. S., Luo, S. B., et al. (2019). A high-density genetic map developed by specific-locus amplified fragment (SLAF) sequencing and identification of a locus controlling anthocyanin pigmentation in stalk of *Zicaitai* (*Brassica rapa* L. ssp. *chinensis* var. *purpurea*). *BMC Genomics* 20, 343. doi: 10.1186/s12864-019-5693-2
- Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191
- Librado, R. J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187
- Liu, L., Fan, X., Tan, P., Wu, J., Zhang, H., Han, C., et al. (2021). The development of SSR markers based on RNA-sequencing and its validation between and within *Carex* L. species. *BMC Plant Biol.* 21, 17. doi: 10.1186/s12870-020-02792-8
- Lohse, M., Drechsel, O., and Bock, R. (2007). OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps

- of plastid and mitochondrial genomes. *Curr. Genet.* 52, 267–274. doi: 10.1007/s00294-007-0161-y
- Ma, Q., Li, S., Bi, C., Hao, Z., Sun, C., and Ye, N. (2017). Complete chloroplast genome sequence of a major economic species, *Ziziphus jujuba* (Rhamnaceae). *Curr. Genet.* 63, 117–129. doi: 10.1007/s00294-016-0612-4
- Ma, Q., Wang, Y., Li, S., Wen, J., Zhu, L., Yan, K., et al. (2022). Assembly and comparative analysis of the first complete mitochondrial genome of *Acer truncatum* Bunge: a woody oil-tree species producing nervonic acid. *BMC Plant Biol.* 22, 29. doi: 10.1186/s12870-021-03416-5
- Martin, W., Stoebe, B., Goremykin, V., Hansmann, S., Hasegawa, M., and Kowallik, K. V. (1998). Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* 393, 162–165. doi: 10.1038/30234
- Møller, I. M., Rasmuson, A. G., and Van Aken, O. (2021). Plant mitochondria - past, present and future. *Plant J.* 108, 912–959. doi: 10.1111/tpj.15495
- Mower, J. (2005). PREP-Mt: predictive RNA editor for plant mitochondrial genes. *BMC Bioinf.* 6, 96. doi: 10.1186/1471-2105-6-96
- Nistor, M., Pop, R., Daescu, A., Pintea, A., Socaciu, C., and Rugina, D. (2022). Anthocyanins as key phytochemicals acting for the prevention of metabolic diseases: an overview. *Molecules* 27, 4254. doi: 10.3390/molecules27134254
- Pinard, D., Myburg, A. A., and Mizrahi, E. (2019). The plastid and mitochondrial genomes of *Eucalyptus grandis*. *BMC Genomics* 20, 132. doi: 10.1186/s12864-019-5444-4
- Podsędek, A. (2017). Natural antioxidants and antioxidant capacity of Brassica vegetables: a review. *LWT - Food Sci. Tech.* 40, 1–11. doi: 10.1016/j.lwt.2005.07.023
- Rice, D., Alverson, A., Richardson, A., Young, G., Sanchez-Puerta, M., Munzinger, J., et al. (2013). Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. *Science* 342, 1468–1473. doi: 10.1126/science.1246275
- Shao, D., Ma, Y., Li, X., Ga, S., and Ren, Y. (2021). The sequence structure and phylogenetic analysis by complete mitochondrial genome of kohlrabi (*Brassica oleracea* var. *gongylodes* L.). *Mitochondrial DNA B Resour.* 6, 2714–2716. doi: 10.1080/23802359.2021.1966341
- Smith, D., Crosby, K., and Lee, R. (2011). Correlation between nuclear plastid DNA abundance and plastid number supports the limited transfer window hypothesis. *Genome Biol. Evol.* 3, 365–371. doi: 10.1093/gbe/evr001
- Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* 13, e1005595. doi: 10.1371/journal.pcbi.1005595
- Wu, J., Liang, J., Lin, R., Cai, X., Zhang, L., Guo, X., et al. (2022). Investigation of Brassica and its relative genomes in the post-genomics era. *Hortic. Res.* 9, uhac182. doi: 10.1093/hr/uhac182
- Wu, Z., Hu, K., Yan, M., Song, L., Wen, J., Ma, C., et al. (2019). Mitochondrial genome and transcriptome analysis of five alloplasmic male-sterile lines in *Brassica juncea*. *BMC Genomics* 20, 348. doi: 10.1186/s12864-019-5721-2
- Ye, N., Wang, X., Li, J., and Bi, C. (2017). Assembly and comparative analysis of complete mitochondrial genome sequence of an economic plant *Salix suchowensis*. *PeerJ* 5, e3148. doi: 10.7717/peerj.3148
- Zhang, C., Ma, H., Sanchez-Puerta, M. V., Li, L., Xiao, J., Liu, Z., et al. (2020). Horizontal gene transfer has impacted *cox1* gene evolution in *Cassia filiformis*. *J. Mol. Evol.* 88, 361–371. doi: 10.1007/s00239-020-09937-1
- Zhang, L., Li, X., Chang, L., Wang, T., Liang, J., Lin, R., et al. (2022). Expanding the genetic variation of *Brassica juncea* by introgression of the *Brassica rapa* genome. *Hortic. Res.* 9, uhab054. doi: 10.1093/hr/uhab054
- Zhang, N., and Jing, P. (2022). Corinths in Brassicaceae: composition, stability, bioavailability, and potential health benefits. *Crit. Rev. Food Sci. Nutr.* 62, 2205–2220. doi: 10.1080/10408398.2020.1852170
- Zhang, X., Zhang, K., Wu, J., Guo, N., Liang, J., Wang, X., et al. (2020). QTL-seq and sequence assembly rapidly mapped the gene *brMYBL2.1* for the purple trait in *Brassica rapa*. *Sci. Rep.* 10, 2328. doi: 10.1038/s41598-020-58916-5
- Zhao, N., Grover, C., Chen, Z., Wendel, J., and Hua, J. (2019). Intergenomic gene transfer in diploid and allopolyploid *Gossypium*. *BMC Plant Biol.* 19, 492. doi: 10.1186/s12870-019-2041-2
- Zhao, N., Wang, Y., and Hua, J. (2018). The roles of mitochondrion in intergenomic gene transfer in plants: a source and a pool. *Int. J. Mol. Sci.* 19, E547. doi: 10.3390/ijms19020547