

RESEARCH

Open Access



Assembly and comparative analysis of the first complete mitochondrial genome of *Acer truncatum* Bunge: a woody oil-tree species producing nervonic acid

Qiuyue Ma¹, Yuxiao Wang², Shushun Li¹, Jing Wen¹, Lu Zhu¹, Kunyuan Yan¹, Yiming Du¹, Jie Ren³, Shuxian Li², Zhu Chen³, Changwei Bi^{2*} and Qianzhong Li^{1*}

Abstract

Background: *Acer truncatum* (purpleblow maple) is a woody tree species that produces seeds with high levels of valuable fatty acids (especially nervonic acid). The species is admired as a landscape plant with high developmental prospects and scientific research value. The *A. truncatum* chloroplast genome has recently been reported; however, the mitochondrial genome (mitogenome) is still unexplored.

Results: We characterized the *A. truncatum* mitogenome, which was assembled using reads from PacBio and Illumina sequencing platforms, performed a comparative analysis against different species of *Acer*. The circular mitogenome of *A. truncatum* has a length of 791,052 bp, with a base composition of 27.11% A, 27.21% T, 22.79% G, and 22.89% C. The *A. truncatum* mitogenome contains 62 genes, including 35 protein-coding genes, 23 tRNA genes and 4 rRNA genes. We also examined codon usage, sequence repeats, RNA editing and selective pressure in the *A. truncatum* mitogenome. To determine the evolutionary and taxonomic status of *A. truncatum*, we conducted a phylogenetic analysis based on the mitogenomes of *A. truncatum* and 25 other taxa. In addition, the gene migration from chloroplast and nuclear genomes to the mitogenome were analyzed. Finally, we developed a novel *NAD1* intron indel marker for distinguishing several *Acer* species.

Conclusions: In this study, we assembled and annotated the mitogenome of *A. truncatum*, a woody oil-tree species producing nervonic acid. The results of our analyses provide comprehensive information on the *A. truncatum* mitogenome, which would facilitate evolutionary research and molecular barcoding in *Acer*.

Keywords: *Acer truncatum*, Mitochondrial genome, Repeats, Phylogenetic analysis

Background

Acer truncatum Bunge (Sapindaceae) is a versatile, oil-producing woody tree widely distributed mainly in northern China, Japan and Korea [1, 2]. This tree species

is a potential source of medicinal compounds, including flavonoids, alkaloids, tannins, and terpenoids [3]. Moreover, *A. truncatum* seed oil contains approximately 90% unsaturated fatty acids and was listed as a new food resource by the Ministry of Health of the People's Republic of China in 2011 [2]. Nervonic acid (24:15, *cis*-15-tetracosenoic acid, n-9) accounts for 5–6% of seed oil [2, 4]. It is a key component of brain nerve cells as well as tissues promoting the repair and regeneration of nerve cells

*Correspondence: bichwei@163.com; qianzhongli@jaas.ac.cn

¹ Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China

² Nanjing Forestry University, Nanjing 210037, China

Full list of author information is available at the end of the article



and damaged tissues. Previous studies have indicated that nervonic acid is potentially useful for treatment of schizophrenia, psychosis, and attention deficit disorder [5, 6]. It has been detected in several plant species [2, 7, 8], but issues related to their nervonic acid content and growth adaptability have limited the utility of these species. The characteristics of rapid growth, wide geographic distribution, and high adaptability, thus *A. truncatum* is a novel potential plant source of nervonic acid for treating human cerebral and neurological problems.

The main function of mitochondria, the “energy factories” of cells, is the conversion of biomass energy into chemical energy in living cells [9, 10]. In most seed plants, nuclear hereditary information is inherited biparentally, whereas DNA of both mitochondria and chloroplasts is maternally derived [9, 10]. In addition, recent researches have revealed that intergenomic gene transfer between nuclear and organellar genomes, which was a common phenomenon during plant evolution [11–13]. Along with rapid developments in sequencing and genome assembly technologies, an increasing amount of information on mitogenomes has been uncovered. At present, 6026 complete land plant organelle genomes, including 5735 chloroplast and 291 plant mitogenomes have been assembled and deposited in GenBank Organelle Genome Resources (<https://www.ncbi.nlm.nih.gov/genome/browse/>), as the mitochondrial genome is more complex and harder to assemble than that of other organelles [9, 14].

Plant mitogenomes are species specific [15, 16] and vary considerably in length, gene order, and gene content [9, 10, 14, 17]. Genome size is extremely variable, ranging from 66 kb (*Viscum scurruloideum*) [18] to 11.3 Mb (*Silene conica*) [19], and most genomes are 200–800 kb in size [20]. This wide variation in mitogenome size can be attributed to the repetitive sequences and the foreign DNA from other organisms during evolution [21, 22]. Repetitive sequences, including simple sequence repeats (SSRs), tandem repeats and dispersed repeats, are abundant in the mitogenomes of seed plants. SSRs are frequently used as molecular markers for identifying species in plant mitogenomes [14, 23]. In addition, insertions/deletions (indels) and single nucleotide polymorphisms (SNPs) within mitogenomes also have been applied to rapidly distinguish species and for phylogenetic analyses [24, 25].

The mitochondrial gene content of land plants varies considerably, ranging from 32 to 67 genes. Some genes, including those related to NADH dehydrogenase, ATP synthase, ubiquinol cytochrome, and cytochrome c biogenesis [14], are highly conserved, whereas others, such as *sdh3*, *sdh4*, *rps11*, and *cox2* have been lost [26, 27].

Mitogenomes in the genus *Acer*, except for the mitogenome sequence of *A. yangbiense* released in 2019, have

not been analyzed in detail [28]. In this study, we first assembled the complete mitogenome of *A. truncatum* and analyzed its gene content, repetitive sequences, RNA editing sites, selective pressure, and phylogenetic relationships. We also surveyed gene transfer among nuclear, chloroplast, and mitochondrial genomes of *A. truncatum*. Moreover, we developed a marker based on an indel in the *NAD1* intron to distinguish seven *Acer* species (*A. buergerianum*, *A. truncatum*, *A. henryi*, *A. negundo*, *A. ginnala*, *A. yangbiense* and *A. tonkinense*). The data presented herein expand genetic information available for the genus *Acer* and provide an opportunity to conduct further important genomic breeding studies on *A. truncatum*.

Results

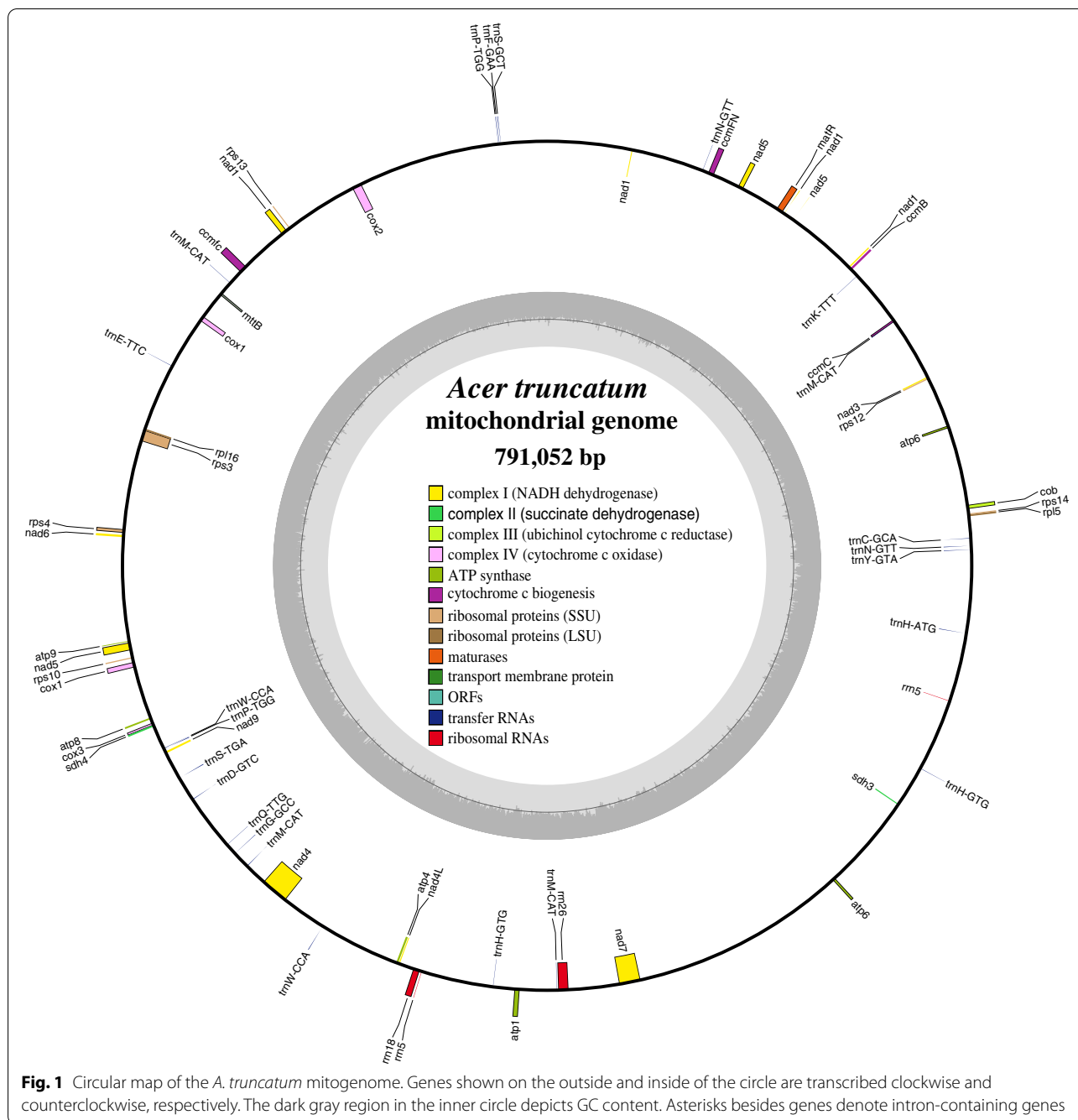
Features of the *A. truncatum* mitogenome

The *A. truncatum* genome sequence generated was submitted to the GenBank database (accession number MZ318049) in this study. The complete mitogenome of *A. truncatum* is 791,052 bp in length and has the typical circular structure of land plant genomes (Fig. 1). The nucleotide composition of the complete mitogenome is 27.11% A, 27.21% T, 22.79% G, and 22.89% C, with a GC content of 45.68% (Table 1). Protein-coding genes (PCGs) and *cis* introns account for 4.31 and 2.94% of the whole mitogenome, while tRNA and rRNA genes comprise only 0.22 and 0.67%, respectively. A total of 62 unique genes, including 35 protein-coding, 23 tRNA, and 4 rRNA genes, were identified in the *A. truncatum* mitogenome (Table 2). Interestingly, two copies of *cox1* genes were found. Additionally, five tRNA and one rRNA gene(s) located in repeat sequences were found to be present in two or four copies (*trnN-GTT*, *trnM-CAT*, *trnP-TGG*, *trnH-GTG*, *trnW-CCA*, and *rrn5*) (Fig. 1).

Codon usage analysis of PCGs

The total length of PCGs in *A. truncatum* was 34,059 bp. Most PCGs had the typical ATG start codon, whereas *atp6*, *nad1*, and *nad4L* had ACG as the start codon—presumably a consequence of C-to-U RNA editing of the second site (Table 2). Three types of stop codons were identified, namely, TAA, TGA, and TAG, the C to U RNA editing phenomenon was not found in the stop codons. As shown in Fig. 2, the codon usage analysis revealed the most frequent amino acids to be leucine (Leu) (11.2–11.3%), serine (Ser) (10.6–11.0%), and arginase (Arg) (8.1–8.4%), whereas cysteine (Cys) and tryptophan (Trp) were rarely found.

We also analyzed the relative synonymous codon usage (RSCU) of 35 PCGs in the *A. truncatum* mitogenome. As shown in Fig. 3, the 35 PCGs comprised



33,948bp encoding 11,316 codons excluding termination codons. We found that nearly all of the RSCU values of NNT and NNA codons were higher than 1.0 with the exception of Ile (AUA, 0.82), Leu (CUA, 0.93), and Ser (UCA, 0.97). Codon usage was generally strongly biased toward A or T(U) at the third codon position in the *A. truncatum* mitogenome, which is very common in mitogenomes of land plant species.

Analysis of synonymous and nonsynonymous substitution rates

In genetics, the nonsynonymous-to-synonymous substitution ratio (Ka/Ks) is used to understand the evolutionary dynamics of genes. In this study, the Ka/Ks ratio was determined for 26 protein-coding genes common to *A. truncatum*, *A. yangbiense*, *A. thaliana* and *C. sinensis* mitogenomes (Fig. 4). The PCGs shared between *A. truncatum* and *A. yangbiense* were close homologs, as the $Ka/$

Table 1 Genomic features of the *A. truncatum* mitogenome

Feature	A(%)	C(%)	G(%)	T(%)	GC(%)	Size (bp)	Proportion in Genome (%)
Whole genome	27.11	22.89	22.79	27.21	45.68	791,052	100
Protein-coding genes	26.12	31.1	21.52	21.25	52.62	34,059	4.31
cis-spliced introns	23.59	26.56	26.56	24.76	53.11	23,222	2.94
tRNA genes	24.59	26.27	24.77	24.36	51.04	1728	0.22
rRNA genes	23.37	26.34	25.74	24.55	52.08	5280	0.67
Non-coding regions	27.30	22.36	22.36	27.60	44.71	72,6763	91.87

Ks ratio of 21 PCGs was 0. In addition, nearly all *Ka/Ks* ratios were less than 1.0, which suggested that most of the PCGs were subject to stabilizing selection during evolution. Conversely, the *Ka/Ks* ratios of nine genes (*atp6*, *cob*, *cox1*, *nad2*, *ccmFn*, *nad4*, *nad6*, *nad7* and *rpl5*) were greater than 1.0, which indicated these genes had been under positive selection during evolution. Finally, three genes (*atp4*, *ccmB* and *rps4*) had *Ka/Ks* ratios close to 1, thus suggested that they had experienced neutral evolution since the divergence of their common ancestor.

Prediction of RNA editing sites in PCGs

In plants, RNA editing is necessary for gene expression, with cytidine (C)-to-uridine (U) RNA editing enriched in mitochondrial and chloroplast genomes. In this study, we predicted the RNA editing sites of 26 PCGs common to mitogenomes of four angiosperm species. The number of RNA editing sites predicted for *A. truncatum*, *A. yangbiense*, *A. thaliana*, and *C. sinensis*—421, 427, 342 and 288, respectively—suggests that these sites are extremely conserved in PCGs in *Acer*. A total of 421 RNA editing sites were predicted in *A. truncatum*, all exhibiting C-to-U RNA editing. Among the 421 sites, 32.07 and 67.93% were predicted at the first and the second positions of codons, respectively, whereas none were found at the third position (Fig. 5).

RNA editing can change PCG initiation and termination codons. As shown in Table 2, *atp6*, *nad1* and *nad4L* genes use ACG as their initiation codons, we infer that they may have been altered by RNA editing. The number of RNA editing sites in different genes was found to vary greatly, with the largest predicted numbers detected in cytochrome *c* biogenesis (*ccmB*, *ccmC*, *ccmFn*, and *ccmFc*), Complex I (NADH dehydrogenase) and *nad4* genes. In contrast, no RNA editing sites were found in *atp9* and *nad3* genes in *A. truncatum* and *A. yangbiense*.

Analysis of repeats in the *A. truncatum* mitogenome

An analysis of repeats in the *A. truncatum* mitogenome revealed 503 long repeats (> 30 bp), namely, 287 forward

(57.05%), 179 palindromic (35.59%), 33 reverse (6.60%) and 1 complementary (0.20%) repeats (Fig. 6A). The total length of the long repeats was 144,318 bp, which corresponded to 18.24% of the mitogenome. Most repeats were 35–50 bp long (254 repeats, 50.29%), whereas 24 were longer than 1 kb, the largest was 28,452 bp (Fig. 6B and Table S1). In *A. truncatum* mitogenome, we found that five pair of large repeats (> 1 kb) by rearrangements could produce two subgenomic circles, which comprising of 457,840 bp and 333,212 bp, mediated by the pairwise large repeats R3a and R3b (Table S2 and Fig.S1). We also identified repeats in the *A. yangbiense* mitogenome to further characterize repeats in *Acer* species. A total of 500 long repeats were found, including 271 forward (54.20%), 88 palindromic (17.60%), and 141 reverse (28.20%) repeats. No complementary repeats were identified. The total length constituted by long repeats was 138,024 bp, which accounted for 17.18% of the *A. yangbiense* mitogenome (803, 281 bp) (Fig. S2A and B). Most repeats were 41–60 bp long (288 repeats, 57.60%), the longest repeat was 27,124 bp (Table S3).

SSRs, which are tandem repeated sequences with motifs of one to six bases, are useful molecular markers for studying genetic diversity and identifying species [14, 23]. In this study, a total of 717 SSRs were detected in the *A. truncatum* mitogenome, including 226 (31.52%) mono-, 355 (49.51%) di-, 49 (6.83%) tri-, 67 (9.34%) tetra-, 18 (2.51%) penta-, and 2 (0.28%) hexanucleotide repeats (Table 3). Among the 717 SSRs, more than 81% were mono- and di-repeats. Further analysis of SSR repeat units indicated that 85.40% of monomers had A/T contents, and 45.07% of dinucleotide repeats were AT/TA. The higher AT content of SSRs contributed to the AT richness (54.32%) of the complete *A. truncatum* mitogenome.

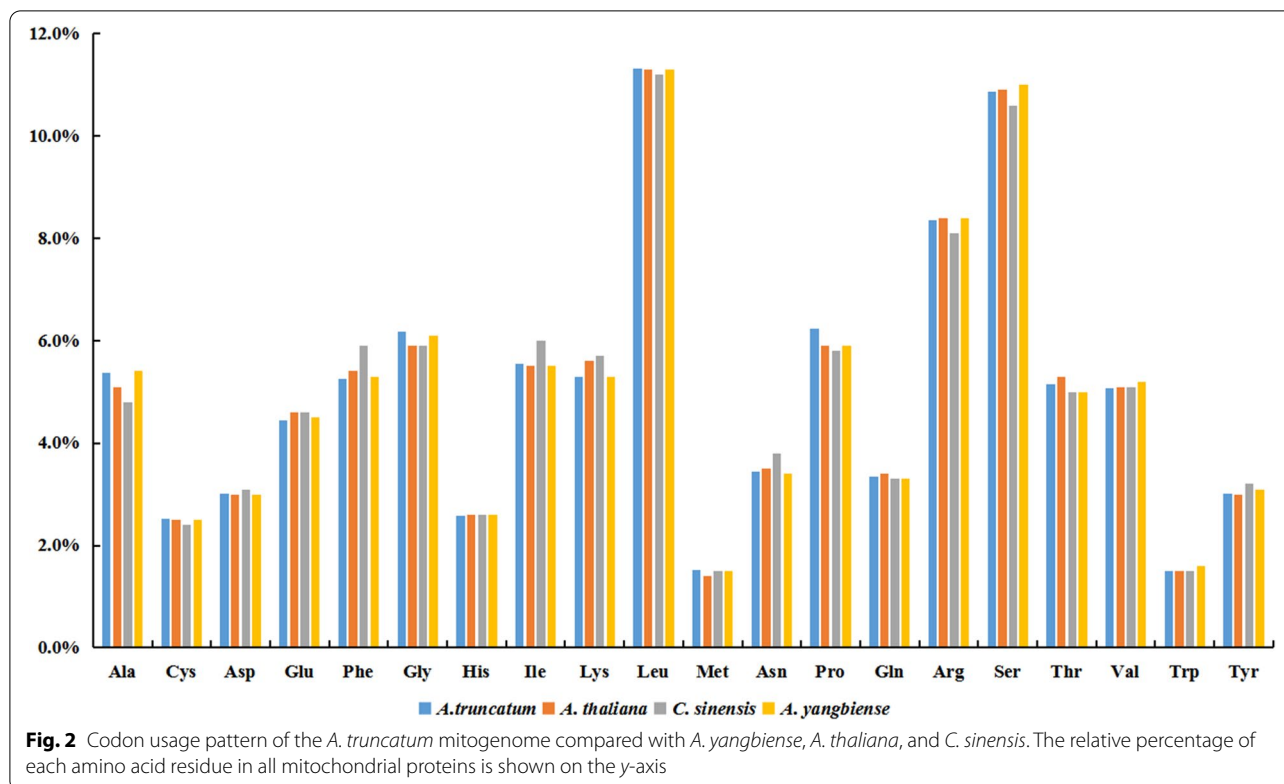
Phylogenetic analysis

To determine the phylogenetic position of *A. truncatum*, we downloaded 25 plant mitogenomes from GenBank (<https://www.ncbi.nlm.nih.gov/genome/browse/>) (Table

Table 2 Gene profile and organization of the *A. truncatum* mitogenome

Group of genes	Gene name	Length	Start codon	Stop codon	Amino acids
ATP synthase	<i>atp1</i>	1530	ATG	TGA	509
	<i>atp4</i>	597	ATG	TAG	198
	<i>atp6</i>	774	ACG	TAA	257
	<i>atp8</i>	480	ATG	TAA	159
	<i>atp9</i>	225	ATG	TGA	74
NADH dehydrogenase	<i>nad1^a</i>	978	ACG	TAA	325
	<i>nad2^a</i>	1467	ATG	TAA	488
	<i>nad3</i>	357	ATG	TAA	118
	<i>nad4^a</i>	1488	ATG	TGA	495
	<i>nad4L</i>	303	ACT	TAA	100
	<i>nad5</i>	2004	ATG	TAA	667
	<i>nad6</i>	618	ATG	TAA	205
	<i>nad7^a</i>	1185	ATG	TAG	394
	<i>nad9</i>	573	ATG	TAA	190
Cytochrome c biogenesis	<i>ccmB</i>	621	ATG	TGA	206
	<i>ccmC</i>	753	ATG	TGA	250
	<i>ccmFc^a</i>	1365	ATG	TAG	454
	<i>ccmFn</i>	1734	ATG	TGA	577
Maturases	<i>matR</i>	1962	ATG	TAG	653
Ubichinol cytochrome c reductase	<i>cob</i>	1182	ATG	TGA	393
Cytochrome c oxidase	<i>cox1(2)</i>	1584	ATG	TAA	527
	<i>cox2</i>	795	ATG	TGA	264
	<i>cox3</i>	798	ATG	TGA	265
Transport membrane protein	<i>mttB</i>	792	ATA	TAG	264
Ribosomal proteins (LSU)	<i>rpl5</i>	555	ATG	TAA	184
	<i>rpl16</i>	516	ATG	TAA	171
Ribosomal proteins (SSU)	<i>rps3^a</i>	1686	ATG	TAA	561
	<i>rps4</i>	1077	ATG	TAA	358
	<i>rps10^a</i>	330	ATG	TAA	109
	<i>rps12</i>	378	ATG	TGA	125
	<i>rps13</i>	294	ATG	TGA	97
	<i>rps14</i>	255	ATG	TGA	84
Succinate dehydrogenase	<i>sdh3</i>	327	ATG	TGA	108
	<i>sdh4</i>	480	ATG	TAA	159
Transfer RNAs	<i>trnY-GTA</i>	83	–	–	–
	<i>trnN-GTT^b(2)</i>	72	–	–	–
	<i>trnC-GCA</i>	71	–	–	–
	<i>trnM-CAT(4)</i>	73/74/74/77	–	–	–
	<i>trnK-TTT</i>	73	–	–	–
	<i>trnS-GCT</i>	88	–	–	–
	<i>trnF-GAA</i>	74	–	–	–
	<i>trnP-TGG^b(2)</i>	74/75	–	–	–
	<i>trnE-TTC</i>	72	–	–	–
	<i>trnW-CCA^b(2)</i>	73/74	–	–	–
	<i>trnS-TGA</i>	87	–	–	–
	<i>trnD-GTC^b</i>	74	–	–	–
	<i>trnQ-TTG</i>	72	–	–	–
	<i>trnG-GCC</i>	72	–	–	–
	<i>trnH-GTG^b(2)</i>	74/74	–	–	–
<i>trnH-ATG</i>	76	–	–	–	
Ribosomal RNAs	<i>rrn5(2)</i>	119/120	–	–	–
	<i>rrn18</i>	1939	–	–	–
	<i>rrn26</i>	3102	–	–	–

Note: Numbers after gene names are the number of copies. The superscripts a and b indicate genes containing introns and chloroplast-derived genes, respectively



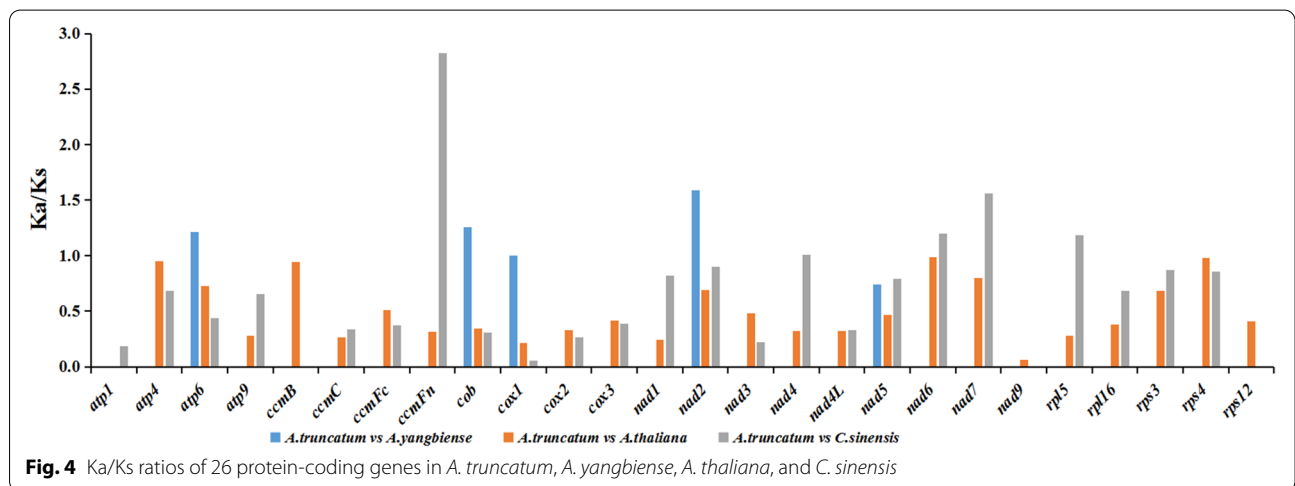
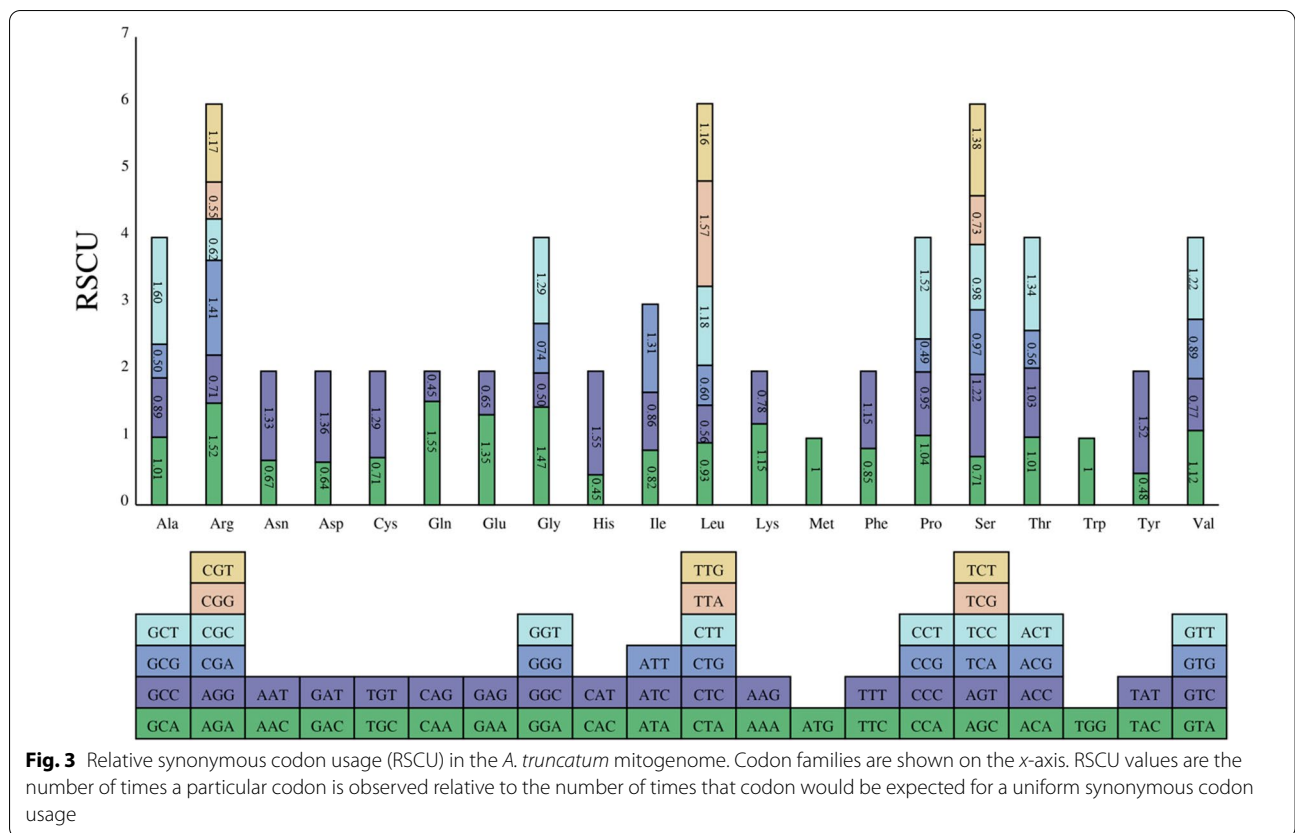
S4) and constructed a phylogenetic tree based on a set of 25 conserved single-copy orthologous genes present in all 26 analyzed mitogenomes. As shown in Fig. 7, 21 of 23 nodes in the generated tree had bootstrap support values over 70%, including 12 nodes with 100% support. The phylogenetic tree strongly supports (100% bootstrap support) the close phylogenetic relationship between *A. truncatum* and *A. yangbiense*. In addition, it also revealed that both species were closely related to *C. sinensis*, which is similar to conclusions inferred using the nuclear genome [2]. Overall, the results of our analysis of mitogenomes provide a valuable foundation for future analyses of the phylogenetic affinities of *Acer* species.

Plastid-derived and nuclear-shared sequence transfer events

DNA fragment transfers among nuclear and organellar genomes are common events during plant evolution. Six directions of gene transfer are possible among the three types of genomes. To further understand the characteristics of sequence transfer events in *A. truncatum*, the *A. truncatum* nuclear and chloroplast genomes [2, 29] were searched by using its mitogenome sequences as queries. We obtained 393 hits covering 230.0kb of sequences of nuclear genome transferred into the mitogenome. According to the nuclear-mitochondrial alignment, hits occurred

on every *A. truncatum* chromosome (Fig. 8A), however, the total lengths of the hits and the percent coverage on the chromosomes were different. Chromosome 1 had the maximum total length of hits (25.30kb), which was much larger than on other chromosomes, whereas the highest percent coverage (0.05%) occurred on chromosomes 5, 6, and 13. In addition, fragment lengths were mainly between 200bp and 400bp (Fig. 8B). A total of 62,241 bp of sequences (7.87% of the *A. truncatum* mitogenome) were found to be shared between nuclear and mitochondrial genomes. The shared sequences contained seven complete genes (*trnN-GTT*, *rpl5*, *trnS-GCT*, *trnF-GAA*, *trnQ-TTG*, *atp1*, and *trnH-GTG*) as well as partial gene sequences of *matR*, *ccmFN*, *cox2*, *rps3*, *rps4*, *atp8*, *sdh4*, *nad4* and *atp6*.

The *A. truncatum* mitogenome sequence (791,052 bp) was approximately five times longer than the chloroplast genome (156,492 bp). Forty-one fragments with a total length of 18,637 bp, corresponding to 2.36% of the mitogenome, were observed to have migrated from the chloroplast genome to the mitogenome in *A. truncatum* (Table 4). Six intact chloroplast genes (*psbJ*, *trnP-UGG*, *trnW-CCA*, *trnN-GUU*, *trnD-GUC*, and *trnH-GUG*) were located on these fragments. The remaining fragments were partial sequences of transferred genes or intergenic spacer regions in the chloroplast genome. Interestingly, we found



that the DNA migration had often occurred in the inverted repeat region of the *A. truncatum* chloroplast genome.

Development of an NAD1 intron indel marker

Among *Acer* species, only the mitogenome of *A. yangbiense* has currently been reported. To further characterize the *NAD1* intron, we compared its sequence

between *A. truncatum* and *A. yangbiense*, and detected a 33-bp indel. The following seven *Acer* species were selected for characterization of the *NAD1* intron sequence: *A. truncatum*, *A. buergerianum*, *A. ginnala*, *A. yangbiense*, *A. palmatum*, *A. pubipalmatum*, and *A. tonkinense*. To develop indel markers, primers were designed to anneal to conserved regions

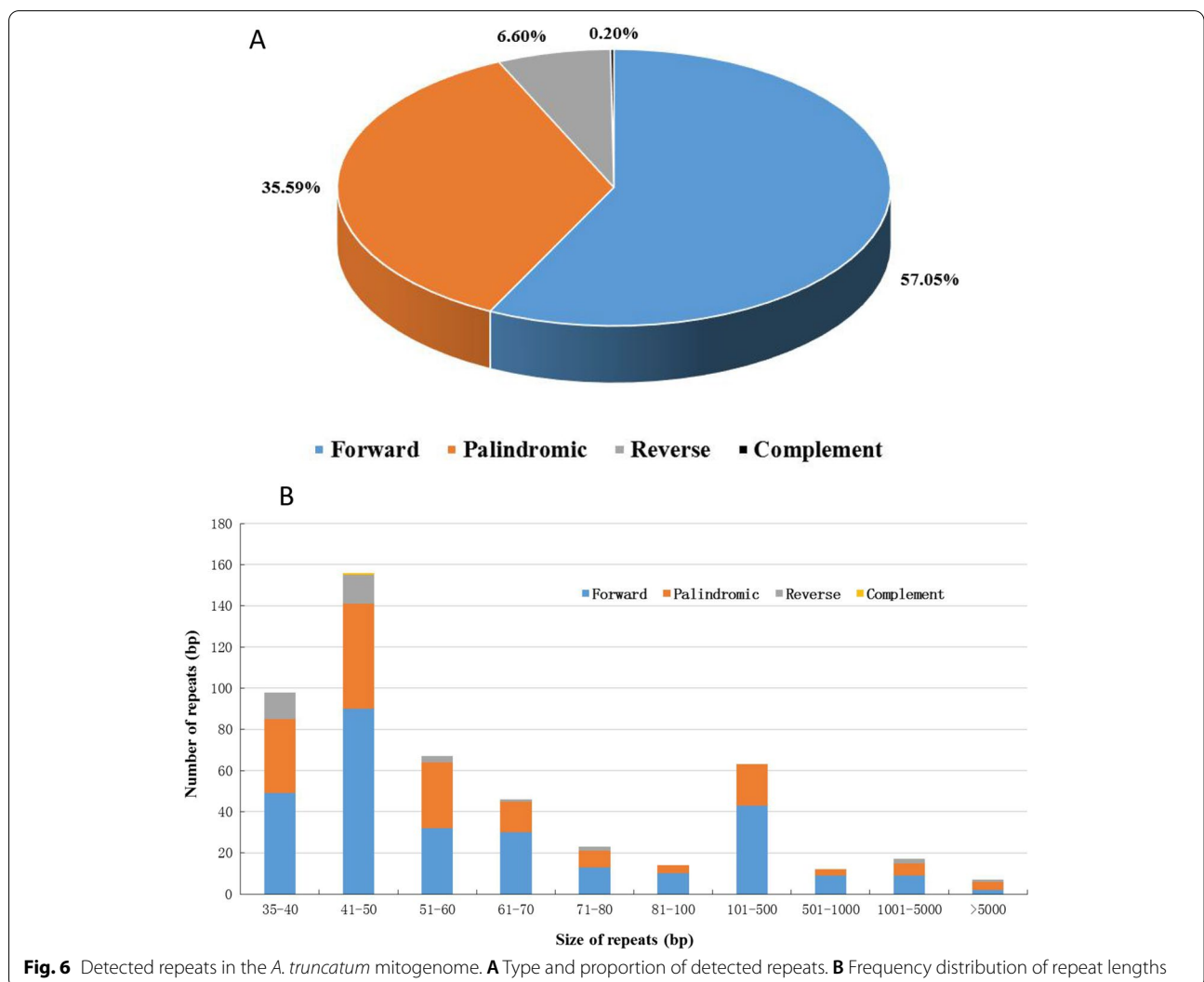
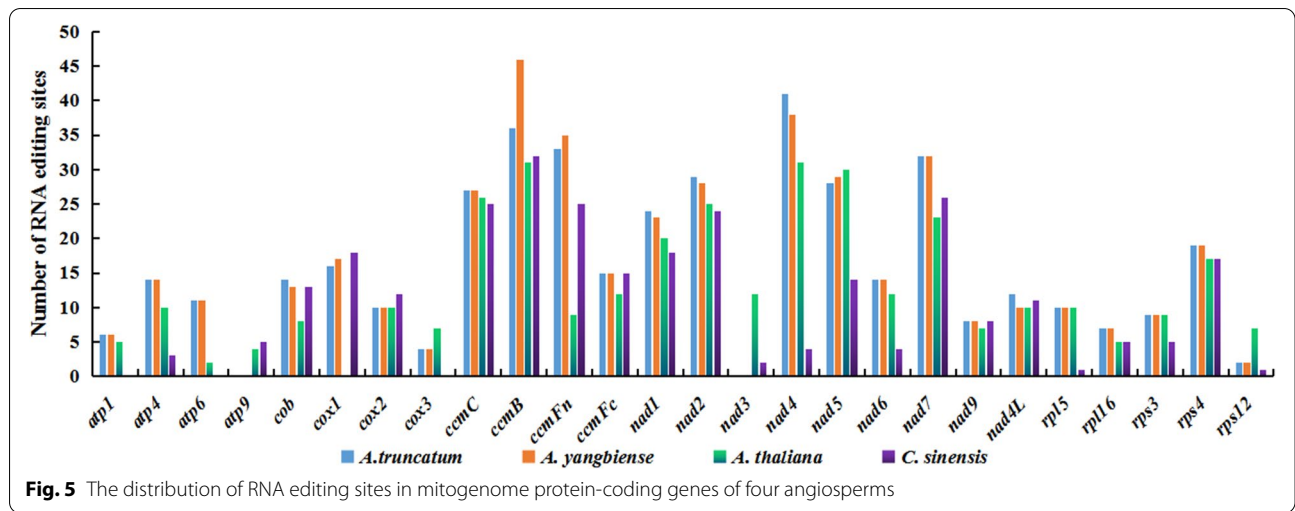
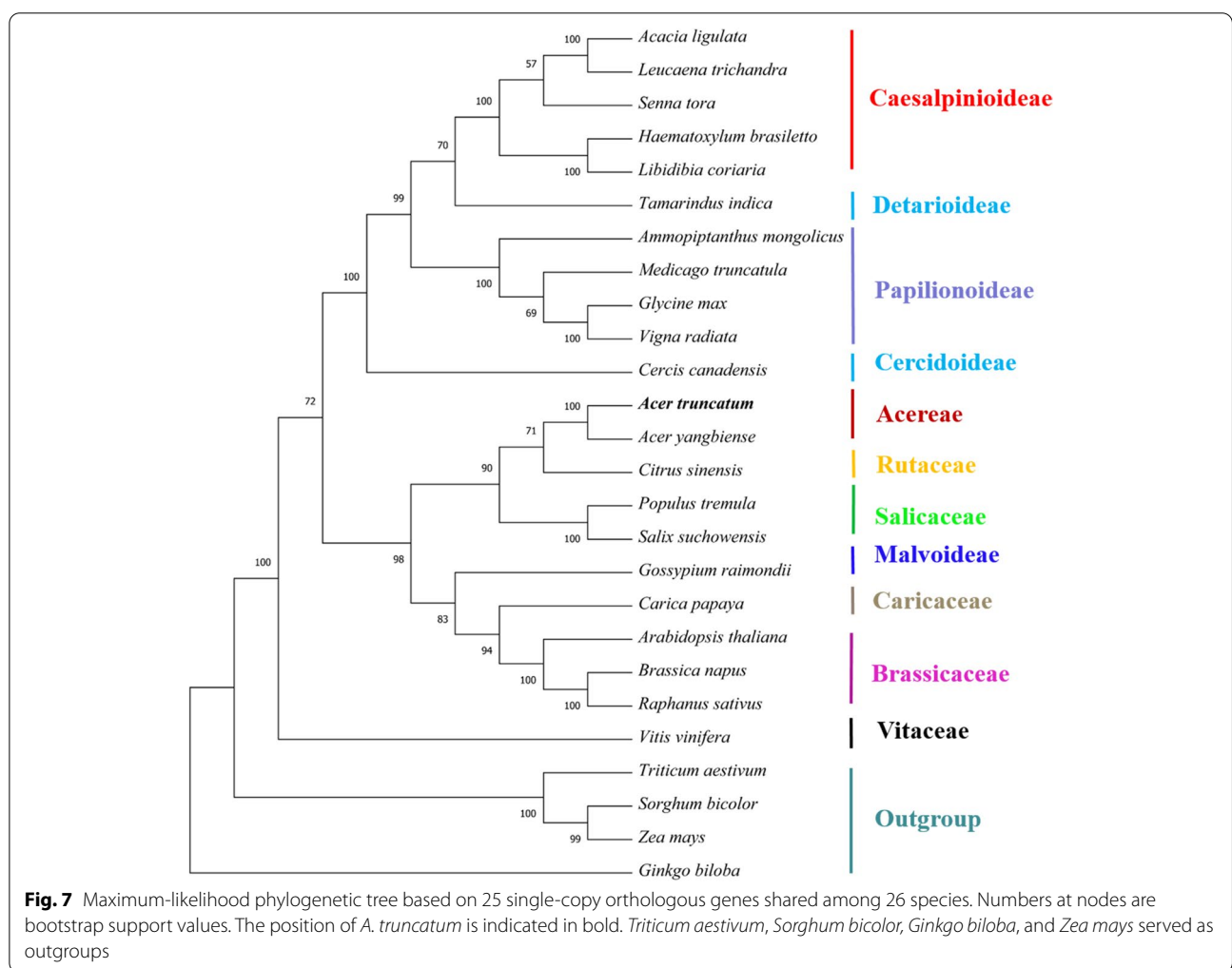


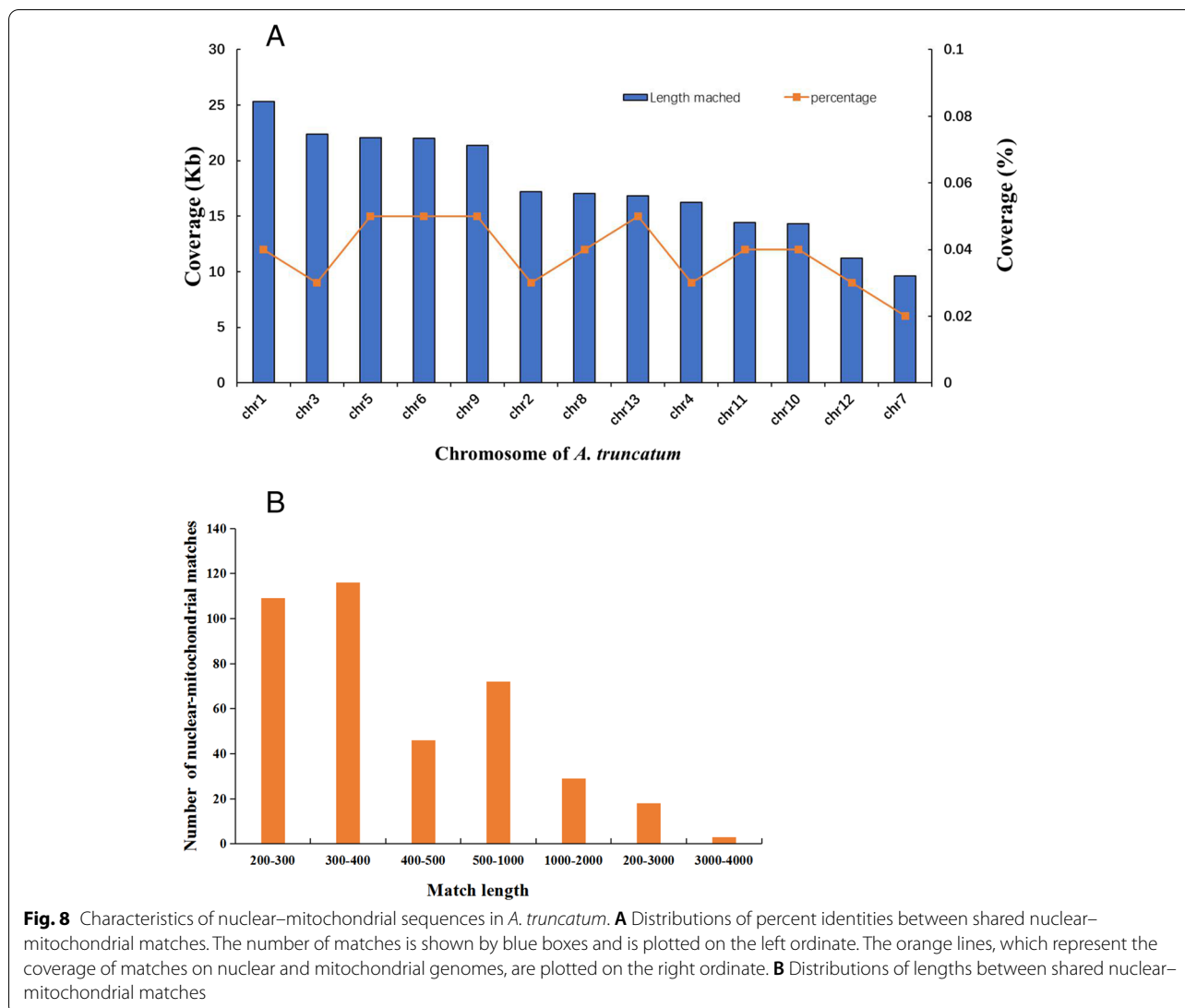
Table 3 Frequency of identified SSR motifs in the *A. truncatum* mitogenome

Motif Type	Number of repeats													Total	Proportion (%)
	3	4	5	6	7	8	9	10	11	12	14	15	21		
Monomer	–	–	–	–	–	118	65	28	7	3	2	2	1	226	31.52
Dimer	–	280	51	14	5	2	2	1	–	–	–	–	–	355	49.51
Trimer	–	43	3	1	1	–	1	–	–	–	–	–	–	49	6.83
Tetramer	59	7	1	–	–	–	–	–	–	–	–	–	–	67	9.34
Pentamer	16	–	–	2	–	–	–	–	–	–	–	–	–	18	2.51
Hexamer	2	–	–	–	–	–	–	–	–	–	–	–	–	2	0.28
Total	77	330	55	17	6	120	68	29	7	3	2	2	1	717	100



of the *NAD1* intron (Table S5). The predicted amplification products were successfully obtained using these *NAD1*-intron-F/R primers in all seven tested samples (Fig. 9A). In all six species, the length of the amplified *NAD1* intron sequence was identical

(808 bp) and highly conserved. The corresponding sequence in *A. yangbiense* was indeed longer (841 bp) because of the 33-bp putative insertion (Fig. 9A and B). Several species close to *A. truncatum* in the phylogenetic tree (*A. yangbiense*, *Populus tremula*, *Salix suchowensis* and *C. sinensis*) were selected to verify



whether the 33-bp sequence was an insertion or a deletion. According to the sequence alignment, the sequence was indeed an insertion (Fig. S3). In previous studies, indel markers have frequently been used to distinguish closely related species; however, *Acer* species have not been identified on the basis of their mitogenomes using this approach. Our first-ever characterization of the *NADI* intron in *Acer* may therefore be applicable for classification and identification of *Acer* species.

Discussion

Characterization of the *A. truncatum* mitochondrial genome

Mitochondria, which produce the energy required to carry out life processes, are the powerhouses of plants. Because of factors such as size variation and repeated

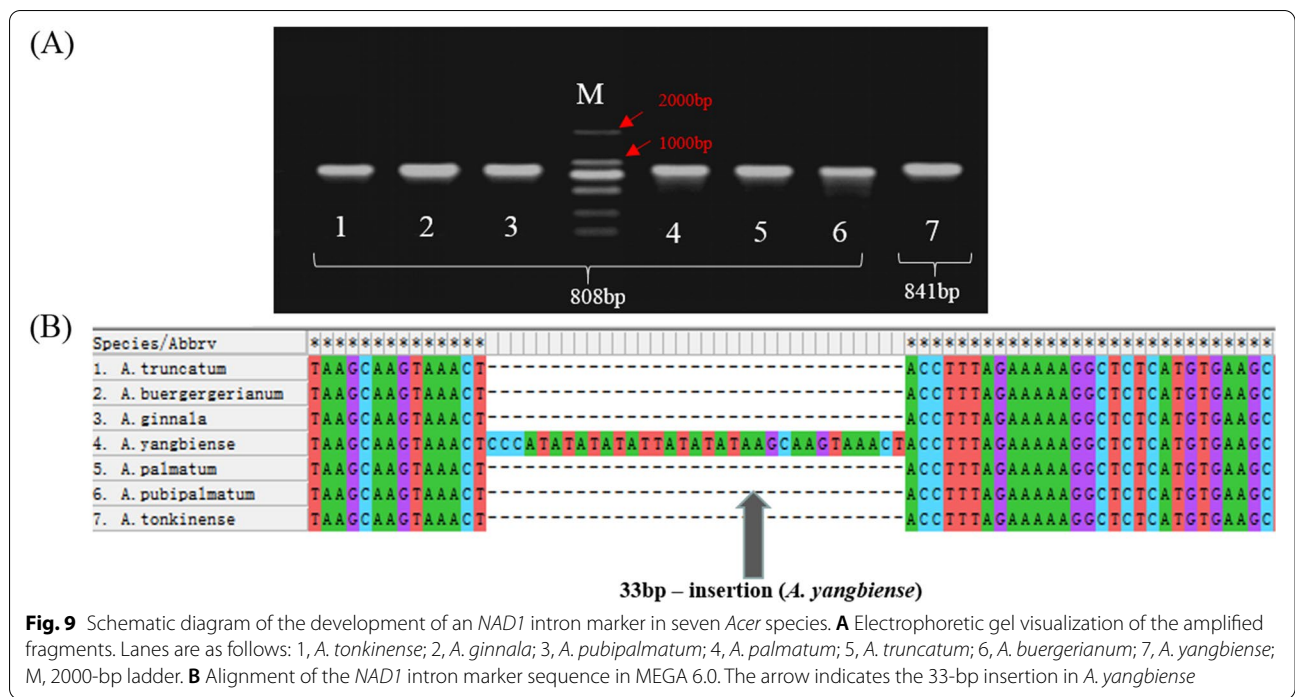
sequences, plant mitogenomes are more complex than those of animals [9, 14, 30, 31]. The emergence of rapid, cost-effective genome sequencing technologies has accelerated understanding of mitogenomes. Our study has produced the first detailed characterization of a complete mitogenome in *Acer*. The size of the *A. truncatum* mitogenome is similar to that of *A. yangbiense* [28], both of which are moderate in size relative to most genomes [32]. GC content is an important factor for assessing species. The GC content of the *A. truncatum* mitogenome is 45.68%, which is comparable to that of other sequenced plant mitogenomes (*A. thaliana*, 44.8% [33]; *Phaseolus vulgaris*, 45.11% [14]; *Beta vulgaris*, 43.9% [34], but higher than the *A. truncatum* chloroplast genome (37.90%) assembled by our research group [29]. Similar to most other mitogenomes, most sequences in the *A. truncatum* mitogenome are non-coding. Protein-coding

Table 4 Fragments transferred from chloroplasts to mitochondria in *A. truncatum*

	Alignment Length	Identity %	Mismatch	Gap opens	CP Start	CP End	Mt Start	Mt End	Gene
1	2890	99.689	9	0	22,796	25,685	588,995	586,106	<i>rpoC1</i>
2	2890	99.239	21	1	22,796	25,685	762,940	760,052	<i>rpoC1</i>
3	2700	99.963	1	0	20,067	22,766	699,997	697,298	
4	1259	97.935	7	8	99,134	100,376	576,186	574,931	
5	1259	97.935	7	8	142,127	143,369	574,931	576,186	
6	1259	97.935	6	9	142,127	143,369	730,678	729,424	
7	1259	97.935	6	9	99,134	100,376	729,424	730,678	
8	1067	90.909	55	28	65,889	66,926	449,842	450,895	<i>psbJ</i>
9	351	99.715	1	0	45,259	45,609	437,233	437,583	<i>ycf3</i>
10	349	99.713	1	0	75,224	75,572	46,646	46,994	<i>psbB</i>
11	224	99.107	1	1	138,184	138,407	589,548	589,326	<i>trnI-GAU</i>
12	224	99.107	1	1	104,096	104,319	589,326	589,548	<i>trnI-GAU</i>
13	205	93.171	12	2	66,231	66,434	362,782	362,579	<i>psbF</i>
14	173	94.798	9	0	68,227	68,399	451,672	451,844	<i>trnP-UGG</i>
15	141	100	0	0	110,171	110,311	150,810	150,670	
16	141	100	0	0	132,192	132,332	150,670	150,810	
17	131	100	0	0	35,802	35,932	277,631	277,761	<i>psbC</i>
18	123	92.683	9	0	67,987	68,109	451,461	451,583	<i>trnW-CCA</i>
19	93	100	0	0	136,588	136,680	694,605	694,513	<i>trnA-UGC</i>
20	93	100	0	0	105,823	105,915	694,513	694,605	<i>trnA-UGC</i>
21	98	97.959	2	0	110,389	110,486	150,666	150,569	
22	98	97.959	2	0	132,017	132,114	150,569	150,666	<i>trnN-GUU</i>
23	178	84.831	17	8	30,780	30,955	468,861	469,030	<i>trnD-GUC</i>
24	105	96.19	3	1	109,900	110,003	533,208	533,104	
25	105	96.19	3	1	132,500	132,603	533,104	533,208	
26	90	98.889	1	0	151,306	151,395	452,466	452,377	<i>ycf2</i>
27	90	98.889	1	0	91,108	91,197	452,377	452,466	<i>ycf2</i>
28	83	100	0	0	59,800	59,882	72,900	72,982	<i>accD</i>
29	86	96.512	3	0	132,009	132,094	5982	5897	<i>trnN-GUU</i>
30	86	96.512	3	0	110,409	110,494	5897	5982	<i>trnN-GUU</i>
31	79	98.734	1	0	54,466	54,544	494,502	494,424	<i>trnM-CAU</i>
32	75	100	0	0	154,389	154,463	77,164	77,090	<i>trnI-CAU</i>
33	75	100	0	0	88,040	88,114	77,090	77,164	<i>trnI-CAU</i>
34	80	97.5	2	0	7	86	577,140	577,219	<i>trnH-GUG</i>
35	80	97.5	2	0	7	86	728,470	728,391	<i>trnH-GUG</i>
36	72	100	0	0	101,734	101,805	206,176	206,105	
37	72	100	0	0	140,698	140,769	206,105	206,176	
38	77	93.506	4	1	7420	7495	131,293	131,217	
39	56	100	0	0	87,421	87,476	346,975	346,920	<i>rpl2</i>
40	56	100	0	0	155,027	155,082	346,920	346,975	<i>rps12(exon)</i>
41	65	92.308	2	3	54,754	54,816	448,383	448,446	<i>atpE</i>
Total	18,637								

genes account for only 4.31%, which is probably the result of a gradual increase in sequence duplication during evolution. Most PCGs were the typical ATG start codon, and the distribution of amino acid compositions was similar to other angiosperms [28, 33]. while the *atp6*, *nad1* and

nad4L genes use ACG as initiation codons, this phenomenon also has been found in other studies, which were considered to be altered by RNA editing modification [9, 14]. The *cox1* is one of the most reported mitochondrial genes involved in horizontal gene transfer among



sngiosperms. In our study, two copies of *cox1* genes were found. The previous studies reported that the *cox1* copies existed in different species and different populations of a species [35]. Ka/Ks ratios > 1 have also been reported for some other mitochondrial genes [9, 14, 26]. In our study, the high Ka/Ks ratios of genes observed were very important for further studies in the gene selection and evolution of *Acer* species, including *atp6*, *cob*, *cox1*, *nad2*, *ccmFn* etc.

Identification of repeat sequences and RNA editing sites

Repeats are important sources of information for developing markers for population and evolutionary analyses [23, 36, 37]. Including tandem, short and large repeats, they are widely present in mitogenomes [14, 38, 39]. Repeats in mitochondrial DNA are generally vital for intermolecular recombination, which can generate structural variations and extreme mitogenome sizes [20, 40]. In this study, five pair of large repeats (>1 kb) by rearrangements could produce two subgenomic circles in *A. truncatum* mitogenome, comprising of 457,840bp and 333,212bp, respectively. This phenomenon also reported in Soybean [41]. we also found major differences between the repeat sequences of *A. truncatum* and *A. yangbiense* mitogenomes. In particular, the proportion of long repeat sequences in the *A. truncatum* mitogenome (18.24%) was higher than that of *A. yangbiense* (17.20%), and the

longest repeats were 28,452bp and 27,124bp, respectively. These repeats may have contributed to the increase in the mitogenome size of *A. yangbiense*. This finding also suggests that intermolecular recombination has frequently occurred in the mitogenome during *Acer* evolution [14, 31].

RNA editing, a post-transcriptional process that occurs in chloroplast and mitochondrial genomes of higher plants, contributes to improved protein folding [9, 14, 26]. Previous researches had uncovered approximately 491 RNA editing sites within 34 genes in rice [42] and 486 RNA editing sites within 31 genes in *P. vulgaris* [14]. In the present study, we predicted RNA editing sites in 26 PCGs common to *A. truncatum*, *A. yangbiense*, *A. thaliana* and *C. sinensis* mitogenomes. We found that the number of RNA editing sites in PCGs was extremely conserved in *Acer* but differed in the other two species. Although the number of RNA editing sites varies greatly among genes, cytochrome *c* biogenesis and NADH dehydrogenase genes harbor the largest number, which is similar to *P. vulgaris* [14]. In addition, all identified RNA editing sites are located at first and second codon positions. Previous researchers have speculated that the lack of RNA editing sites at the third codon position is probably due to the limitations of the PREP-Mt predictive methodology used rather than an actual absence [14, 43]. Further analysis using experimental methods is thus needed.

DNA fragment transfer events

Information pertaining to DNA transfer events between different genomes (mitochondrial, nuclear and chloroplast) has been uncovered by sequencing analysis [21, 44, 45]. Previous studies have determined that the most prominent transfer direction in angiosperms is from organellar genomes into the nuclear genome, followed in importance by transfer from nuclear and plastid genomes into the mitogenome [13, 21, 46–48]. The total length of transferred DNA varies among plant species in higher plants, lengths range from 50 kb (*A. thaliana*) to 1.1 Mb (*O. sativa* subsp. *japonica*) [49]. According to our study, 230.0 kb of nuclear DNA has been transferred into the mitogenome of *A. truncatum*. Although the nuclear–mitochondrial transfer process has occurred on every *A. truncatum* chromosome, the total lengths of transferred material and the percent coverage differs among chromosomes. In total, 62, 241 bp of sequences (7.87% of the *A. truncatum* mitogenome) is shared between nuclear and mitochondrial genomes. Most genes with transferred sequences shared between nuclear and mitochondrial genomes are tRNA genes, such as *trnN-GTT*, *trnH-GTG*, and *trnH-GTG*. Chang et al. [41] have reported similar results in soybean. In regards to chloroplast genome to mitogenome migration events, a total of 18,637 bp of transferred fragments were observed, accounting for 2.36% of the *A. truncatum* mitogenome. In comparison, the proportion in *S. suchowensis* and *Suaeda glauca* is 2.8 and 5.18%, respectively [31]. We identified 41 fragments that had been transferred from the chloroplast genome to the mitogenome, these fragments included six integrated genes, namely, five tRNA genes and *psbJ*. Transfer of tRNA genes from chloroplast to mitochondrial DNA is common in angiosperms [21, 26, 31]. Interestingly, we also observed that DNA migration often occurred in the inverted repeat region of the *A. truncatum* chloroplast genome.

Development of a mitochondrial *NAD1* intron marker for *Acer* species

Because indel regions are relatively easy to detect, they are often used to develop markers for identifying species [50]. The genus *Acer* comprises more than 200 species grown in China [2, 51]; however, the highly similar shapes of some species present a challenge for identification, and a molecular approach would be beneficial. *NAD1* intron indel markers have been useful for identification of some plant species [52–54]. In *Acer*, only the mitogenome of *A. yangbiense* has been previously reported [28]. In the present study, we first identified a 33-bp sequence difference by aligning the *NAD1* intron regions of *A. truncatum* and *A. yangbiense*. Amplification of the *NAD1* intron with specific primers revealed that

a 33-bp indel was present in *A. yangbiense*, whereas the amplified *NAD1* intron sequence was of the same length and highly conserved in the other six species. We verified that this 33-bp indel was an insertion in *Acer* by analyzing several species close to *A. truncatum* in our phylogenetic tree (*A. yangbiense*, *P. tremula*, *S. suchowensis*, and *C. sinensis*). The development of mitogenome-based molecular markers has not been previously reported for *Acer*. Although only a few *Acer* species were used in this study, our findings should nonetheless contribute to species classification in *Acer*.

Conclusions

In this study, we assembled and annotated the mitogenome of *A. truncatum* and performed extensive analyses based on DNA and amino acid sequences of annotated genes. The *A. truncatum* mitogenome is circular, with a length of 791,052 bp. We annotated 62 genes, including 35 protein-coding, 23 tRNA and 4 rRNA genes. In addition, the codon usage, sequence repeats, RNA editing and selective pressure were also analyzed in the *A. truncatum* mitogenome. The evolutionary status of *A. truncatum* was verified by phylogenetic analysis based on the mitogenomes of this species and 25 other taxa. Gene conservation between chloroplast and mitochondrial genomes and between nuclear and mitochondrial genomes were also detected in *A. truncatum* by analyzing gene migration. Finally, a newly developed *NAD1* intron indel marker was used to distinguish *Acer* species. Our study has yielded extensive information about the *A. truncatum* mitogenome. The data presented herein supplement the genetic knowledge available for the genus *Acer*, provide novel insights into *A. truncatum* evolution, and form an important theoretical basis for increasing *A. truncatum* seed yield.

Materials and methods

Plant materials and DNA sequencing

A. truncatum plants were grown at our Aceraceae seed base of Jiangsu Academy of Agricultural Sciences (Lishui District, Nanjing, China; 31°65'N, 119°02'E) under natural conditions. Fresh leaves were frozen in liquid nitrogen and stored at 80 °C. DNA extraction and sequencing were performed using methods described in our previous de novo genome sequencing study [2].

Mitogenome assembly and annotation

For the *A. truncatum* mitogenome, PacBio RS II reads (59.42 GB) sequenced in our previous study [2] were de novo assembled using Canu v1.4 [55]. The obtained contigs were mapped to core mitochondrial genes by mini-map2 [56], then extended. The assembled contigs were

polished (Pilon v1.18) with Illumina reads (75.0 GB) to correct read errors [57]. Finally, five large contigs were assembled into mitogenome, to verify the quality and accuracy of our assemblies, we further verified the junctions by Sanger sequencing (Table S6 and Fig. S4). The GE-Seq tool on the MPI-MP CHLOROBOX website (<https://chlorobox.mpimp-golm.mpg.de>) was used for the mitogenome annotation, with the *A. yangbiense* mitogenome (CM017774.1) serving as a reference. Mitochondrial protein-coding genes were predicted using the MITOFY webserver [40]. All tRNA genes were confirmed by using tRNAscan-SE with default settings [9, 14, 58]. ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>) was used to analyze open reading frames longer than 300 bp. RSCU values and the amino acid composition of PCGs were calculated in MEGA X [59]. A circular mitochondrial map was drawn using Organellar Genome DRAW [60].

Analysis of repeat structures and SSRs

Forward, reverse, palindromic and complementary repeats were identified with REPuter [61]. SSRs were analyzed with the MISA program [62]. The motif size of one- to six- nucleotide SSRs was set as 8, 4, 4, 3 and 3, respectively.

Selective pressure analysis

We calculated the nonsynonymous (K_a) and synonymous (K_s) substitution rates of each PCG between *A. truncatum* and *A. yangbiense*, *A. thaliana*, and *C. sinensis*. Orthologous gene pairs were separately aligned in MEGA 6.0. K_a , K_s , and K_a/K_s values were calculated using DnaSP [63].

Genome alignments

The *A. truncatum* mitogenome was searched against the chloroplast genome of *A. truncatum* (MH638284) using BLASTN 2.9.0+ according to the following screening criteria: matching rate $\geq 70\%$, E-value $\leq 1e^{-6}$, and length ≥ 40 [31]. To identify regions of potential nuclear origin in the mitogenome of *A. truncatum*, we also performed a BLASTN search (maximum E-value = $1e^{-50}$) of the complete mitogenome against all contigs from the *A. truncatum* nuclear genome sequenced in our previous study. BLASTN results of sequences longer than 250 bp and a pairwise similarity $> 80\%$ were inspected for sequence features.

Prediction of RNA editing sites

RNA editing sites in the PCGs of *A. truncatum* and other three mitogenomes (*A. yangbiense*, *A. thaliana* and *C. sinensis*) were predicted using the the online PREP-Mt

suite of servers (<http://prep.unl.edu/>). To obtain a more accurate prediction, the cutoff value was set as 0.2 [43].

Phylogenetic analyses

A total of 26 complete mitogenomes (Table S4) were used to ascertain the phylogenetic position of *A. truncatum*. The 25 mitochondrial PCG genes (*atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *ccmB*, *ccmC*, *ccmFc*, *ccmFn*, *cob*, *cox1*, *cox3*, *matR*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *nad9*, *rps12*, *rps3*, and *rps4*) conserved across the 26 analyzed species were aligned in Muscle with default parameters [64], with the alignment then modified manually to eliminate gaps and missing data. Finally, a maximum likelihood tree was constructed in MEGA X using the JTT+G+I+F nucleotide substitution model [58]. A bootstrap consensus tree was inferred from 1000 bootstrap replicates. *Triticum aestivum*, *Sorghum bicolor*, *Ginkgo biloba*, and *Zea mays* were used as outgroups.

Verification of the *Nad1* insertion in *Acer*

Primers were designed with Primer 5. PCR amplifications were carried out in 15- μ l volumes containing 20 ng genomic DNA, 0.4 μ l dNTPs (2.5 mM each), 2.5 μ l of 10 \times Ex *Taq* buffer (Mg^{2+}), 0.4 μ l Ex *Taq* DNA polymerase (Takara, Tokyo, Japan), and 1.0 μ l of each primer (10 mM). The amplification conditions were 94 °C for 5 min, followed by 30 cycles of 94 °C for 30 s, 56 °C for 30 s, and 72 °C for 30 s, with a final extension of 72 °C for 10 min. The PCR products were purified and linked to the pMD19-T easy plasmid (Takara) for sequencing to confirm the accuracy of PCR product sizes. Three samples per species were sequenced by the General Biology Company (Nanjing, Jiangsu, China).

Abbreviations

Ile: Isoleucine; Leu: Leucine; Ser: Serine; PCR: Polymerase chain reaction; SSRs: Simple sequence repeat; PCG: Protein-coding genes; DNA: Deoxyribonucleic acid; SNPs: Single Nucleotide Polymorphism; tRNA: Transfer RNA; rRNA: Ribosomal RNA; Arg: Arginine; Cys: Cysteine; Trp: Tryptophan; RSCU: Relative synonymous codon usage; K_a/K_s : Nonsynonymous-to-synonymous substitution ratio.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-021-03416-5>.

Additional file 1: Figure S1. Reversible reorganization of the *A. truncatum* mitogenome may produce subgenomic circles by large repeats. The same colour triangles represent the pairs of large repeats.

Additional file 2: Figure S2. The type of detected repeats and the frequency distribution of lengths in the *A. yangbiense*.

Additional file 3: Figure S3. Alignment of the *NAD1* intron sequence with MEGA-X.

Additional file 4: Figure S4. Agarose gel electrophoresis of PCR product for contig connecting verification

Additional file 5: Table S1. The repeat sequences distributions in the *A. truncatum* mitogenome genome.

Additional file 6: Table S2. The large repeats (> 1 kb) by rearrangements could produce two subgenomic circles in *A. truncatum* mitogenome.

Additional file 7: Table S3. The repeat sequences distributions in the *A. yangbiense* mitogenome genome.

Additional file 8: Table S4. Details regarding the mitochondria genome sequences used for the phylogenetic analysis.

Additional file 9: Table S5. Details regarding the primers used to develop the *NAD1* intron marker.

Additional file 10: Table S6. Primers for contig connecting verification

Acknowledgments

Thanks to all the members of the Institute of Leisure Agriculture team for their contributions, and the Mitogenome assembly and annotation by the Genepioneer Biotechnologies Co, Ltd.

Authors' contributions

QYM, QZL, CWB designed the project and the strategy, SXL, JW, ZC and YMD contributed to plant sample collection; QYM, YXW, CWB, LZ, JR and KYY work on genome assembly, annotation and comparative analyses; QYM, CWB and QZL wrote and revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was funded by the Natural Science Foundation of China (32001357), Natural Science Foundation of Jiangsu Province (BK20211139), the Independent Innovation Fund Project of Agricultural Science and Technology in Jiangsu Province (CX (20)3155).

Availability of data and materials

The *A. truncatum* Mitochondrial genome sequence was deposited in the GenBank database (accession number MZ318049).

Declaration

Competing interests

The authors declare that they have no conflicts of interest.

Author details

¹Institute of Leisure Agriculture, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, China. ²Nanjing Forestry University, Nanjing 210037, China. ³Institute of Agricultural Engineering, Anhui Academy of Agricultural Sciences, 40 Nongkenanlu, Hefei 230031, Anhui, China.

Received: 14 August 2021 Accepted: 27 December 2021

Published online: 13 January 2022

References

- Guo X, Wang R, Chang R, Liang X, Wang C, Luo Y, et al. Effects of nitrogen addition on growth and photosynthetic characteristics of *Acer truncatum* seedlings. *Dendrobiology*. 2014;72:151–61. <https://doi.org/10.12657/denbio.072.013>.
- Ma Q, Sun T, Li S, Wen J, Zhu L, Yin T, et al. The *Acer truncatum* genome provides insights into nervonic acid biosynthesis. *Plant J*. 2020;104(3):662–78. <https://doi.org/10.1111/tpj.14954>.
- Tang W, Wang J, Xu J, Wang L, Huang J, Chen Y. Advances of chemical composition of medicinal plants in Aceraceae. *Northern Horticulture*. 2012;18:194–200.
- Wang X, Wang S. A new resource of nervonic acid from purpleblow maple (*Acer truncatum*) seed oil. *For Prod J*. 2005;09:60–2.
- Tanaka K, Shimizu T, Ohtsuka Y, Yamashiro Y, Oshida K. Early dietary treatments with Lorenzo's oil and docosahexaenoic acid for neurological development in a case with Zellweger syndrome. *Brain and Development*. 2007;29:586–9. <https://doi.org/10.1016/j.braindev.2007.02.005>.
- Amminger G, Schäfer M, Klier C, Slavik J, Holzer I, Holub M. Decreased nervonic acid levels in erythrocyte membranes predict psychosis in help-seeking ultra-high-risk individuals. *Mol Psychiatry*. 2012;17:1150–67. <https://doi.org/10.1038/mp.2011.167>.
- Taylor D, Francis T, Guo Y, Brost J, Katavic V, Mietkiewska E. Molecular cloning and characterization of a *KCS* gene from *Cardamine graeca* and its heterologous expression in *Brassica* oilseeds to engineer high nervonic acid oils for potential medical and industrial use. *Plant Biotechnol J*. 2009;7:925–38. <https://doi.org/10.1111/j.1467-7652.2009.00454.x>.
- Guo Y, Mietkiewska E, Francis T, Katavic V, Brost J, Giblin M. Increase in nervonic acid content in transformed yeast and transgenic plants by introduction of a *Lunaria annua* L. 3-ketoacyl-CoA synthase (*KCS*) gene. *Plant Mol Biol*. 2009;69:565–75. <https://doi.org/10.1007/s11103-008-9439-9>.
- Ye N, Wang X, Li J, Bi C. Assembly and comparative analysis of complete mitochondrial genome sequence of an economic plant *Salix suchowensis*. *Peer J*. 2017;5(1):e3148. <https://doi.org/10.7717/peerj.3148>.
- Birky C. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc Natl Acad Sci*. 1995;92:11331–8.
- Cusimano N, Wicke S. Massive intracellular gene transfer during plastid genome reduction in nongreen Orobanchaceae. *New Phytol*. 2016;210(2):680–93. <https://doi.org/10.1111/nph.13784>.
- Bock R. Witnessing genome evolution: experimental reconstruction of endosymbiotic and horizontal gene transfer. *Annu Rev Genet*. 2017;51(1):1–22. <https://doi.org/10.1146/annurev-genet-120215-035329>.
- Zhao N, Grover C, Chen Z, Wendel J, Hua J. Intergenomic gene transfer in diploid and allopolyploid *Gossypium*. *BMC Plant Biol*. 2019;19(1):492. <https://doi.org/10.1186/s12870-019-2041-2>.
- Bi C, Lu N, Xu Y, He C. Characterization and analysis of the mitochondrial genome of common bean (*Phaseolus vulgaris*) by comparative genomic approaches. *Int J Mol Sci*. 2020;21(11):3778. <https://doi.org/10.3390/ijms21113778>.
- Hsu C, Mullin B. Physical characterization of mitochondrial DNA from cotton. *Plant Mol Biol*. 1989;13:467–8. <https://doi.org/10.1007/BF00015558>.
- Greiner S, Bock R. Tuning a ménage à trois: co-evolution and co-adaptation of nuclear and organellar genomes in plants. *Drug Alcohol Depend*. 2013;35:354–65. <https://doi.org/10.1002/bies.201200137>.
- Richardson A, Rice D, Young G, Alverson A, Palmer J. The "Fossilized" mitochondrial genome of *Liriodendron tulipifera*: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. *BMC Biol*. 2013;11:29. <https://doi.org/10.1186/1741-7007-11-29>.
- Skippington E, Barkman T, Rice D, Palmer J. Miniaturized mitogenome of the parasitic plant *Viscum scurruloideum* is extremely divergent and dynamic and has lost all *nad* genes. *Proc Natl Acad Sci U S A*. 2015;112(27):E3515. <https://doi.org/10.1073/pnas.1504491112>.
- Sloan D, Alverson A, Chuckalovcak J, Wu M, Mccauley D. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol*. 2012;10(1):e1001241. <https://doi.org/10.1371/journal.pbio.1001241>.
- Guo W, Felix G, Fan W, Young G. *Ginkgo* and *Welwitschia* Mitogenomes reveal extreme contrasts in gymnosperm mitochondrial evolution. *Mol Biol Evol*. 2016;33(6):1448–60. <https://doi.org/10.1093/molbev/msw024>.
- Bergthorsson U, Adams K, Thomason B, Palmer J. Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature*. 2003;424(6945):197–201. <https://doi.org/10.1038/nature01743>.
- Wynn E, Christensen A. Repeats of unusual size in plant mitochondrial genomes: identification, incidence and evolution. *G3: genes, genomes, Genetics*. 2019;9:549–59. <https://doi.org/10.1534/g3.118.200948>.
- Ma Q, Li S, Bi C, Hao Z, Sun C, Ye N. Complete chloroplast genome sequence of a major economic species, *Ziziphus jujuba* (Rhamnaceae). *Curr Genet*. 2017;63:117–29. <https://doi.org/10.1007/s00294-016-0612-4>.
- Seok J, Kim A, Wang J, Hee K, Iksoo L. Single-nucleotide polymorphism markers in mitochondrial genomes for identifying *Varroa destructor*-resistant and -susceptible strains of *Apis mellifera* (Hymenoptera: Apidae). *Mitochondrial DNA Part A, DNA Mapp Seq Anal*. 2019;30(3):477–89. <https://doi.org/10.1080/24701394.2018.1551385>.
- Mwamuye M, Obara I, Khawla E, Odongo D, Bakheit M, Jongejan F, et al. Unique mitochondrial single nucleotide polymorphisms demonstrate

- resolution potential to discriminate *Theileria parva* vaccine and Buffalo-derived strains. *Life*. 2020;10:334. <https://doi.org/10.3390/life10120334>.
26. Bi C, Paterson A, Wang X, Xu Y, Wu D, Qu Y, et al. Analysis of the complete mitochondrial genome sequence of the diploid cotton *Gossypium raimondii* by comparative genomics approaches. *Biomed Res Int*. 2016;2016:5040598. <https://doi.org/10.1155/2016/5040598>.
 27. Choi I, Schwarz E, Ruhlmann T, Khiyami M, Sabir J, Hajarrah N, et al. Fluctuations in Fabaceae mitochondrial genome size and content are both ancient and recent. *BMC Plant Biol*. 2019;19(1):448. <https://doi.org/10.1186/s12870-019-2064-8>.
 28. Yang J, Wariss H, Tao L, Zhang R, Yun Q, Hollingsworth P, et al. De novo genome assembly of the endangered *Acer yangbiense*, a plant species with extremely small populations endemic to Yunnan Province, China. *Gigascience*. 2019;8(7):giz085. <https://doi.org/10.1093/gigascience/giz085>.
 29. Ma Q, Wang Y, Zhu L, Bi C, Li S, Li S, et al. Characterization of the complete chloroplast genome of *Acer truncatum* bunge (Sapindales: Aceraceae): a new woody oil tree species producing nervonic acid. *Biomed Res Int*. 2019;11:1–13. <https://doi.org/10.1155/2019/7417239>.
 30. Kozik A, Rowan B, Lavelle D, Berke L, Schranz M, Michelmore R, et al. The alternative reality of plant mitochondrial DNA: one ring does not rule them all. *PLoS Genet*. 2019;15(8):e1008373. <https://doi.org/10.1371/journal.pgen.1008373>.
 31. Cheng Y, He X, Priyadarshani S, Wang Y, Ye L, Qin Y, et al. Assembly and comparative analysis of the complete mitochondrial genome of *Suaeda glauca*. *BMC Genomics*. 2021;22:167. <https://doi.org/10.1186/s12864-021-07490-9>.
 32. Chang S, Yang T, Du T, Huang Y, Chen J, et al. Mitochondrial genome sequencing helps show the evolutionary mechanism of mitochondrial genome formation in *Brassica*. *BMC Genomics*. 2021;12:497. <https://doi.org/10.1186/1471-2164-12-497>.
 33. Sloan D, Wu Z, Sharbrough J. Correction of persistent errors in *Arabidopsis* reference mitochondrial genomes. *Plant Cell*. 2018;30(3):525–7. <https://doi.org/10.1105/tpc.18.00024>.
 34. Kubo T, Nishizawa S, Sugawara A, Itchoda N, Estiati A, Mikami T. The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNACys (GCA). *Nucleic Acids Res*. 2000;28:2571–6. <https://doi.org/10.1093/nar/28.13.2571>.
 35. Zhang C, Ma H, Sanchez-Puerta MV, Li L, Xiao J, et al. Horizontal gene transfer has impacted *cox1* gene evolution in *Cassytha filiformis*. *J Mol Evol*. 2020;88(4):361–71. <https://doi.org/10.1007/s00239-020-09937-1>.
 36. Xiong Y, Lei X, Bai S, Xiong Y, Liu W, Wu W, et al. Genomic survey sequencing, development and characterization of single- and multi-locus genomic SSR markers of *Elymus sibiricus* L. *BMC Plant Biol*. 2021;21:3. <https://doi.org/10.1186/s12870-020-02770-0>.
 37. Liu L, Fan X, Tan P, Wu J, Zhang H, Han C, et al. The development of SSR markers based on RNA-sequencing and its validation between and within *Carex* L. species. *BMC Plant Biol*. 2021;21:17. <https://doi.org/10.1186/s12870-020-02792-8>.
 38. Gualberto J, Milesina D, Wallet C, Niazi A, Weber-Lotfi F, Dietrich A. The plant mitochondrial genome: dynamics and maintenance. *Biochimie*. 2014;100:107–20. <https://doi.org/10.1016/j.biochi.2013.09.016>.
 39. Guo W, Zhu A, Fan W, Mower J. Complete mitochondrial genomes from the ferns *Ophioglossum californicum* and *Ptilotum nudum* are highly repetitive with the largest organellar introns. *New Phytol*. 2017;213(1):391–403. <https://doi.org/10.1111/nph.14135>.
 40. Alverson AJ, et al. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol Biol Evol*. 2010;27(6):1436–48. <https://doi.org/10.1093/molbev/msq029>.
 41. Chang S, Wang Y, Lu J, Gai J, Li J, Chu P, et al. The mitochondrial genome of soybean reveals complex genome structures and gene evolution at intercellular and phylogenetic levels. *PLoS One*. 2031;8(2):e56502. <https://doi.org/10.1371/journal.pone.0056502>.
 42. Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, et al. The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Mol Gen Genomics*. 2002;268(4):434–45. <https://doi.org/10.1007/s00438-002-0767-1>.
 43. Mower J. PREP-Mt: predictive RNA editor for plant mitochondrial genes. *BMC Bioinformatics*. 2005;6:96. <https://doi.org/10.1186/1471-2105-6-96>.
 44. Timmis J, Ayliffe M, Huang C, Martin W. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet*. 2004;5:123–35. <https://doi.org/10.1038/nrg1271>.
 45. Nguyen V, Giang V, Waminal N, Park H, Kim N, Jang W, et al. Comprehensive comparative analysis of chloroplast genomes from seven *Panax* species and development of an authentication system based on species-unique single nucleotide polymorphism markers. *J Ginseng Res*. 2000;44:135–44. <https://doi.org/10.1016/j.jgr.2018.06.003>.
 46. Martin W, Stoebe B, Goremykin V, Hansmann S, Hasegawa M, Kowallik KV. Gene transfer to the nucleus and the evolution of chloroplasts. *Nature*. 1998;393(6681):162–5. <https://doi.org/10.1038/30234>.
 47. Zhao N, Wang Y, Hua J. The roles of mitochondrion in intergenomic gene transfer in plants: a source and a pool. *Int J Mol Sci*. 2018;19(2):E547. <https://doi.org/10.3390/ijms19020547>.
 48. Rice D, Alverson A, Richardson A, Young G, Sanchez-Puerta M, Munzinger J, et al. Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. *Science*. 2013;342(6165):1468–73. <https://doi.org/10.1126/science.1246275>.
 49. Smith D, Crosby K, Lee R. Correlation between nuclear plastid DNA abundance and plastid number supports the limited transfer window hypothesis. *Genome Biol Evol*. 2011;3:365–71. <https://doi.org/10.1093/gbe/evr001>.
 50. Sebbenn A, Blanc-Jolivet C, Mader M, Meyer-Sand B, Degen B. Nuclear and plastidial snp and indel markers for genetic tracking studies of *Jaccharanda copaia*. *Conservation Genet Resour*. 2019;2019(11):341–3. <https://doi.org/10.1007/s12686-019-01097-9>.
 51. Xu T. The systematic evolution and distribution of the genus *Acer*. *Acta Botanica Yunnanica*; 1998.
 52. Chase F, Freudenstein J. Chase analysis of mitochondrial *nad1b-c* intron sequences in Orchidaceae: utility and coding of length-change characters. *Syst Bot*. 2001;26(3):643–57. <https://doi.org/10.1043/0363-6445-26.3.643>.
 53. Zhang T, Wang Z, Xu L, Zhou K. Application of mitochondrial *nad 1* intron 2 sequences to molecular identification of some species of *dendrobium* Sw. *Chinese Traditional and Herbal Drugs*. 2005;36(7):1059–62.
 54. Pan Z, Ren X, Zhao H, Liu L, Tan Z, Qiu F. A mitochondrial transcription termination factor, *zmsmk3*, is required for *nad1* intron4 and *nad4* intron1 splicing and kernel development in maize. *G3: Genes Genomes Genetics*. 2019;9(8):2677–86. <https://doi.org/10.1534/g3.119.400265>.
 55. Koren S, Walenz B, Berlin K, Miller J, Phillippy A, Canu. Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res*. 2017;27:722–36. <https://doi.org/10.1101/gr.215087.116>.
 56. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34(18):3094–100. <https://doi.org/10.1093/bioinformatics/bty191>.
 57. Walker B, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*. 2014;9(11):e112963. <https://doi.org/10.1371/journal.pone.0112963>.
 58. Lowe T, Eddy S. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997;25(5):955–64. <https://doi.org/10.1093/nar/25.5.955>.
 59. Kumar S, Stecher G, Li M, Niyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35(6):1547–9. <https://doi.org/10.1093/molbev/msy096>.
 60. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res*. 2019;47(W1):W59–64. <https://doi.org/10.1101/545509>.
 61. Kurtz S, Choudhuri J, Enno O, Chris S, Jens S, Robert G. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res*. 2001;29(22):4633–42. <https://doi.org/10.1093/nar/29.22.4633>.
 62. Thiel T, Michalek W, Varshney R, Graner A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor Appl Genet*. 2003;106:411–22. <https://doi.org/10.1007/s00122-002-1031-0>.
 63. Librado R. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 2009;25(11):1451–2. <https://doi.org/10.1093/bioinformatics/btp187>.

64. Edgar R. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32(5):1792–7. <https://doi.org/10.1093/nar/gkh340>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

