



HAL
open science

De Novo Hybrid Assembly Unveils Multi-Chromosomal Mitochondrial Genomes in *Ludwigia* Species, Highlighting Genomic Recombination, Gene Transfer, and RNA Editing Events

Guillaume Dore, Dominique D. Barloy, Frédérique Barloy-Hubler

► **To cite this version:**

Guillaume Dore, Dominique D. Barloy, Frédérique Barloy-Hubler. De Novo Hybrid Assembly Unveils Multi-Chromosomal Mitochondrial Genomes in *Ludwigia* Species, Highlighting Genomic Recombination, Gene Transfer, and RNA Editing Events. *International Journal of Molecular Sciences*, 2024, 25 (13), pp.7283. 10.3390/ijms25137283. hal-04649278

HAL Id: hal-04649278

<https://hal.science/hal-04649278v1>

Submitted on 16 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Article

De Novo Hybrid Assembly Unveils Multi-Chromosomal Mitochondrial Genomes in *Ludwigia* Species, Highlighting Genomic Recombination, Gene Transfer, and RNA Editing Events

Guillaume Doré ¹, Dominique Barloy ^{1,*} and Frédérique Barloy-Hubler ^{2,*}

¹ DECOD (Ecosystem Dynamics and Sustainability), Institut Agro, INRAE, IFREMER, 35042 Rennes, France; guillaume.dore@agrocampus-ouest.fr

² UMR 6553 ECOBIO, CNRS, Université de Rennes 1, 35042 Rennes, France

* Correspondence: dominique.barloy@agrocampus-ouest.fr (D.B.); fhubler@univ-rennes1.fr (F.B.-H.)

† These authors contributed equally to this work.

Abstract: Biological invasions have been identified as the fifth cause of biodiversity loss, and their subsequent dispersal represents a major ecological challenge. The aquatic invasive species *Ludwigia grandiflora* subsp. *hexapetala* (*Lgh*) and *Ludwigia peploides* subsp. *montevidensis* (*Lpm*) are largely distributed in aquatic environments in North America and in Europe. However, they also present worrying terrestrial forms that are able to colonize wet meadows. To comprehend the mechanisms of the terrestrial adaptation of *Lgh* and *Lpm*, it is necessary to develop their genomic resources, which are currently poorly documented. We performed de novo assembly of the mitogenomes of *Lgh* and *Lpm* through hybrid assemblies, combining short reads (SR) and/or long reads (LR) before annotating both mitogenomes. We successfully assembled the mitogenomes of *Lgh* and *Lpm* into two circular molecules each, resulting in a combined total length of 711,578 bp and 722,518 bp, respectively. Notably, both the *Lgh* and *Lpm* molecules contained plastome-origin sequences, comprising 7.8% of the mitochondrial genome length. Additionally, we identified recombinations that were mediated by large repeats, suggesting the presence of multiple alternative conformations. In conclusion, our study presents the first high-quality mitogenomes of *Lpm* and *Lgh*, which are the only ones in the Myrtales order found as two circular molecules.

Keywords: water primrose; Myrtales; mitogenome; hybrid assemblies; DNA transfer; repeated sequences



Citation: Doré, G.; Barloy, D.; Barloy-Hubler, F. De Novo Hybrid Assembly Unveils Multi-Chromosomal Mitochondrial Genomes in *Ludwigia* Species, Highlighting Genomic Recombination, Gene Transfer, and RNA Editing Events. *Int. J. Mol. Sci.* **2024**, *25*, 7283. <https://doi.org/10.3390/ijms25137283>

Academic Editor: Abir U. Igamberdiev

Received: 4 June 2024

Revised: 28 June 2024

Accepted: 29 June 2024

Published: 2 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Origin, Biological Traits, and Distribution of the Aquatic Invasive Plants, *Ludwigia* Species

Biological invasions are now recognized as one of the primary drivers behind global biodiversity loss. Invasive alien plant species (IAPs) lead to significant economic and ecological losses by causing disruption to ecosystems [1]. IAPs are characterized by their strong growth, excellent competitiveness, high adaptive value, an absence of predators, and also reproduce abundantly [2,3].

Of all these invasive aquatic plant species, *Ludwigia grandiflora* subsp. *hexapetala* (or *Lgh*) and *Ludwigia peploides* subsp. *montevidensis* (or *Lpm*) stand out in particular. They're native to South America, but both species have colonized many countries around the world [4]. In North America, *Lgh* and *Lpm* are distributed across several states where they have both been found to degrade major watersheds as well as aquatic and riparian ecosystems [5]. *Lpm* and *Lgh* show some advantageous biological traits that might explain their success of colonization in new areas. *Lpm* and *Lgh* reproduce essentially by clonal propagation with high grow rate, which contribute to the dispersal and establishment of propagules [6]. Both

Ludwigia species are capable of thriving in a wide range of environments, revealing their broad ecological tolerance and great plasticity [7]. Sexual production is also effective in both *Ludwigia* species. In France, *Lpm* is self-compatible and produces many capsules and seeds [8]. *Lgh*, on the other hand, possesses a heteromorphic late-acting self-incompatible system and is also able to produce seeds [9]. Seeds can be dispersed by floating fruit, which contributes to long-distance colonization, as is the case for propagules [7]. *Lgh* has provoked similar damage in Europe, and has been observed in France [10], Germany [11], Italy [12], Spain [13], Switzerland [14], Great Britain/Ireland [15], and Turkey [16]. *Lgh* is also present in Japan [17]. *Lpm* is very common in the southeast and west of France, but its distribution is more limited in Belgium, Corsica, Greece, Italy, the Netherlands, Spain, Turkey, the Balkans, and the UK [18,19]. The oldest documented introductions for both species were in France in 1830 [20].

In its natural habitat, both *Ludwigia* ssp. exhibit an amphibious nature [21], and are commonly found in wetland areas or in transitional zones between aquatic and terrestrial environments [22,23]. However, it is now developing particularly worrying terrestrial forms and is able to colonize wet meadows at an alarming rate [24]. *Lgh* has two distinct morphological types (morphotypes) depending on its environment (aquatic or terrestrial). The two morphotypes present different metabolic processes: fermentation for the aquatic morphotype, and respiration for the terrestrial morphotype [25].

1.2. Genomic Resources in the Onagraceae Family

Genomics plays a central role in identifying genes and metabolic pathways involved in the adaptation process leading to plant invasion [3], with genomic information helping to predict and control invasiveness [26]. For instance, in species such as *Reynoutria japonica* and *Phalaris arundinacea*, genes that play a critical role in invasion are associated with stress response, reproduction, and growth [27,28]. Hence, the genomics of invasive plants represents a powerful tool for comprehending the molecular foundations of plant invasion, enabling the prediction and management of their proliferation [29,30].

Lgh and *Lpm* belong to the malvids clade, Myrtales order, Onagraceae family, and Ludwigioideae subfamily. To date, *Ludwigia* species have not undergone comprehensive genome sequencing, with the exception of the chloroplast genomes of *L. octovalvis* [31], *Lgh* and *Lpm* [32]. In the related subfamily Onagroideae, two incomplete draft genomes are available for *Chamaenerion angustifolium* and *Oenothera biennis* (unpublished GenBank records). *Lpm*, one of the progenitor species of *Lgh*, is a diploid species ($2n = 16$) with a relatively small nuclear genome (261.7 Mb), and *Lgh* is a decaploid ($2n = 80$) with a very large nuclear genome (1.4 Gb) [33]. Assembly of the *Lgh* genome is anticipated to be challenging and time-consuming, as has been acknowledged in previous studies dealing with such big and complex genomes [34]. Having successfully assembled and annotated their plastomes [32], we decided to subsequently assemble their mitogenomes due to the crucial role they play in the plants' biological functions [35,36]. In addition, no Ludwigioideae mitogenomes are currently available, and only four Onagraceae mitogenomes: *Chamaenerion angustifolium* (OX328283) and *Oenothera elata*, *O. biennis*, and *O. villaricae* [37].

1.3. Form and Size of Plant Mitogenomes

Commonly depicted as circular, mitogenomes of terrestrial plants exhibit intricate, diverse, and dynamic structures. Hence, certain plants possess multipartite mitochondrial genomes consisting of several chromosomes, predominantly in a circular form, although they can be linear or branched [38–41]. The most commonly reported structure consists of two circular mitochondrial chromosomes, as demonstrated by numerous studies for *Allium cepa* [42], *Saccharum officinarum* [43], *Fallopia multiflora* [44], *Salvia officinalis* [45], and *Zantedeschia odorata* [46]. Another common structure consists of three circular molecules like *Cynanchum wilfordii* [47], *Brassica campestris* [48], and cucumber [49]. Recent studies have revealed the existence of even more intricate mitochondrial genomes. For instance, *Amborella trichopoda* has five circular chromosomes [50], *P. micranthum* has 26 circular

subgenomes [51], *Lophophytum mirabile* is divided into 54 circular chromosomes [52], *Silene noctiflora* has 59 to 63 circular chromosomes [53], and *Silene conica* has at least 128 circular chromosomes, some of which are empty (i.e., contain no functional genes) [54].

In addition to these structural variations, angiosperm mitochondrial genomes are particularly variable in size. While the mitogenome of *Viscum scurruloideum* is frequently referenced in articles as being the smallest, measuring 66 kb [55], it is worth noting that smaller unpublished mitogenomes exist in the GenBank database including that of Asteraceae *Carthamus tinctorius* (63 kb, OQ621746) or Orchidaceae *Dendrobium amplum* (48 kb, MH591889). In contrast, the largest mitogenome currently described is that of *Larix sibirica* with a total of 11.7 Mb [56].

1.4. Repeat Sequences and Exogenous DNA in Plant Mitogenomes

Vascular plant mitogenomes also contain a large fraction of repeat sequences, making up between five and ten percent of the genome [57]. Repetitive elements, especially within the Angiosperm clade, can reach significant sizes, with repeats exceeding 10 kb being relatively common [58]. A portion of these repeated sequences can be traced back to transposable elements such as retrotransposons and mitoviruses [59,60], and numerous TE fragments have been reported in plant mitogenomes [61–63]. The presence of these extensive repeated sequences within mitogenomes has made it difficult to assemble them accurately [57,64]. Repeats, especially those exceeding 1000 bp, play a significant role in the homologous recombination phenomena [65,66], resulting in different conformations coexisting within the same species [67,68].

The ability of the mitochondrial genome to incorporate exogenous DNA also partly explains these size variations [69]. The most frequent transfer concerns the insertion of chloroplast DNA sequences [70], although sequences of nuclear origin can also be found inserted [71]. Nucleotide sequences derived from plastids and nuclei represent 1–12% and 0.1–13.4% of the mitogenome, respectively [71,72]. Most of these insertions occur in the intergenic regions, even though substitutions of protein-coding gene fractions have been cited [73]. One of the roles of chloroplast-derived sequences is to provide tRNA genes for mitochondrial translation [74], although these sequences may also function as promoters for mitochondrial genes [75]. Most of these chloroplast foreign sequences are non-coding or pseudogenic, although they sometimes appear in coding regions [76] due in particular to alterations in the genetic code between the two organelles [77] and/or differences in the initiation and termination signals of transcription and translation [78]. Remarkably, the amounts of DNA attributed to external sources (plastids and nuclei) are insufficient to account for the origins of all the intergenic regions because more than 60% of these regions show no detectable homology with other known sequences, and their conservation is too weak to support strong functional constraints [79].

1.5. Limits and Objectives of This Study

Another striking feature of angiosperm mitochondrial genomes is their low gene density and a very low collinearity [80,81]. Thus, although the size of their plant mitogenomes is variable, their conserved genetic pool is generally small, with sequences encoding central mitochondrial functions such as subunits of oxidative phosphorylation chain complexes (OXPHOS), proteins involved in the biogenesis of these complexes, and several ribosomal proteins [70,82]. Mitochondrial protein-coding genes typically undergo RNA modification processes to fulfill their functional roles through RNA editing, a post-transcriptional regulation mechanism [83,84].

Due to this remarkable plasticity, it is difficult to assemble plant mitochondrial genomes [85]. However, the recent hybrid assembly approach, which combines long and short reads has proven to be suitable [38,86,87].

In order to better understand the terrestrial adaptation mechanisms of *Lgh* and *Lpm*, and to augment the genomic reservoir of both species, we carried out de novo assemblies of the mitogenomes of *Lgh* and *Lpm*. To achieve this, we first generated *Lpm* and *Lgh* Illumina

Mi-seq short-read (SR) sequences and/or Oxford Nanopore Technologies long reads (LR). Next, we used hybrid assemblies to combine the SR and LR sequences, facilitating the assembly of *Lgh* and *Lpm* mitogenomes. Afterward, we completed functional annotations and analyzed repeats and chloroplast insertions to determine the *Lpm* and *Lgh* mitogenome structures. Finally, we identified genes under selection and constructed a phylogenetic tree using mitogenomes from the Myrtales order.

2. Results

2.1. De Novo Assembly of *Ludwigia* Multichromosomal Mitogenomes

Sequencing of *Lgh* via Oxford Nanopore Technologies yielded a total of 482,619 reads (denoted as LRs for long reads), amounting to 2,386,362,939 bp. These reads exhibited an average length of 4944 bp, with the longest read extending to 72.9 kb. The sequencing conducted with Illumina Mi-seq technology resulted in $2 \times 23,067,490$ reads (abbreviated as SRs for short reads), totaling 9,606,411,104 bp. The reads had an average length of 208 base pairs, with the maximum read length reaching 301 base pairs. Error correction using Ratatosk for LRs and SPAdes for SRs retained 96% of the short reads and all long reads. Using these corrected reads, we conducted three distinct de novo assemblies using Megahit for SR-only assembly, Flye for LR-only assembly, and a hybrid assembly combining both read types using SPAdes.

Statistics for *Lgh* mitogenome assemblies were calculated using QUAST and are summarized in Table S1. We were able to observe that the SPAdes assembly stood out for generating the largest contig (192 kb), the greatest total length (806 kb), and achieving the highest genome fraction (94.9%). However, it also exhibited the highest error rate, with 147 mismatches per 100 kb and three misassemblies. Conversely, Megahit exhibited the lowest error rate, tallying just 46 mismatches per 100 kb and zero misassembled contigs. However, it showed weakness in terms of completion, and only reconstructed 89% of the genome, which stood out as the lowest completion rate among the three assemblers. Finally, Flye emerged as the optimal choice, striking a balance between the two other assemblers, with no misassemblies and less mismatches (71 per 100 kb) than SPAdes, and better LGA50 and total aligned length (664 kb) than Megahit. Simultaneously, with these assemblies, we selected 1118 mitochondrial LRs using coding sequences (CDS) from the malvids clade species mitogenomes. These reads were assembled into 12 contigs with the Geneious assembler, allowing us to identify mitochondrial contigs from the de novo assemblies. Subsequently, we extended all of these contigs to achieve the complete mitogenome. Ultimately, we successfully assembled the *Lgh* mitogenome into two circular molecules, referred to as M1 and M2, with respective sizes of 544,782 bp and 166,796 bp, resulting in a combined total of 711,578 bp (Figure 1). The *Lgh* mitogenome coverage depths were 38x for LRs and 229x for SRs, and these values were similar for both M1 and M2. The *Lgh* mitogenome GC content was 45.2%.

Sequencing *Lpm* using Oxford Nanopore Technologies generated a dataset comprising 68,907 reads, which was reduced to 67,384 following self-correction by RATATOSK. These cumulatively amounted to 510,468,382 bp. These reads exhibited an average length of 7434 bp, with some extending as long as 83.03 kilobases. Without the availability of short reads, only the Flye assembly could be carried out and generated 496 contigs. Using the *Lgh* mitogenome as a reference, we found one *Lpm* contig of 158,656 kb that corresponded to 95% of the *Lgh* M2 molecule, and eight *Lpm* contigs that covered 83.02% of the *Lgh* M1 molecule (453,279 kb in total). Using these contigs and incorporating *Lpm* LRs that were not assembled by Flye, we successfully reconstructed the *Lpm* mitogenome, forming two circular molecules (also designated as M1 and M2), mirroring the structure seen in *Lgh*. These two molecules had respective sizes of 555,518 bp and 167,000 bp, culminating in a total mitogenome size of 722,518 bp (Figure 1). The coverage depth of the *Lpm* mitogenome was 10.1 for M1 and 5.8 for M2, with a GC content of 45.2%.

The mitogenomes of *Lgh* and *Lpm* were submitted to GenBank under the accession numbers PP727126, PP72717, PP727128, and PP727129.

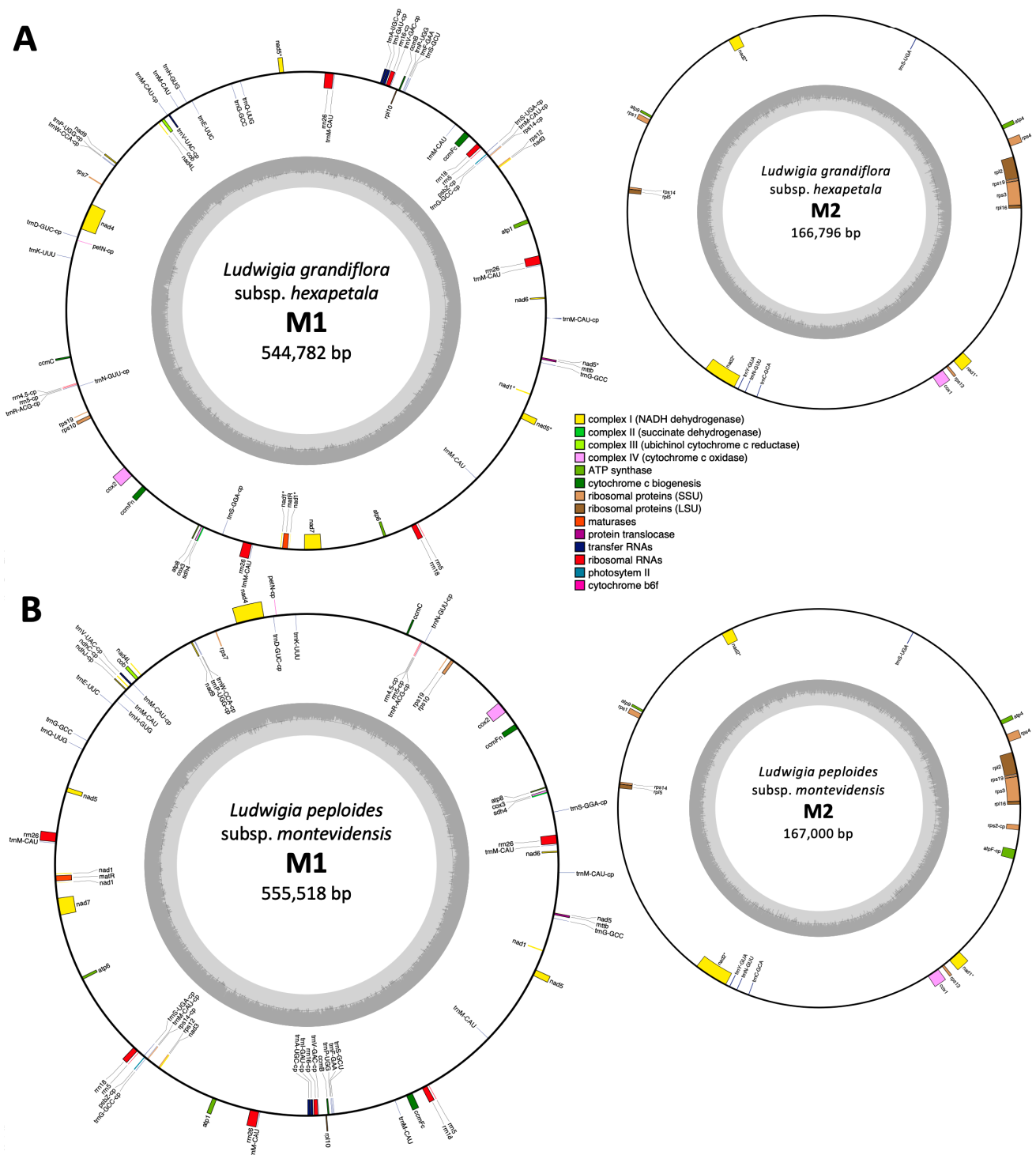


Figure 1. The two circular mitochondrial genomes of *Ludwigia grandiflora* subsp. *hexapetala* (A) and *Ludwigia peploides* subsp. *montevidensis* (B). Genes outside the circle are transcribed in a counter-clockwise direction, while inner genes are transcribed clockwise. Spliced genes are marked with an asterisk *. Genes labeled with -cp originate from the chloroplast. Inner gray circles represent the GC content. Genes are color-coded, as indicated in the legend.

2.2. Annotation and Comparisons of *Ludwigia* Mitogenome Contents

2.2.1. Mitogenome Organization Comparisons

To assess the rearrangement and collinearity of the *Lgh* and *Lpm* mitogenomes, their mtDNA sequences were compared using ProgressiveMauve and Lastz (Figure 2). While

the two M2 molecules were colinear and identical to 99.5%, the M1 counterparts were divided into eight locally collinear blocks (LCBs), collectively representing 88.43% and 90.27% of their total lengths. The sizes of these LCBs were well-preserved, with that of *Lpm* being consistently slightly larger than that of *Lgh* (Table S2). The LCB positions varied, with (as shown in Figure 2C) three inverted blocks following a central symmetry (LCB3, LCB4, LCB7), highlighting extensive mitogenome rearrangements between the two species of *Ludwigia*. In the five inter-LCB regions larger than 500 bp, three contained a repeated sequence of 5253 bp that was conserved between the two *Ludwigia* species. Two of these contained the genes *rrn26* and *trnM-CAU*, and the third contained a region of approximately 6156 kb, also conserved between the two species, carrying the genes *rrn18* and *rrn5*, of which a copy was present at the end of LCB3.

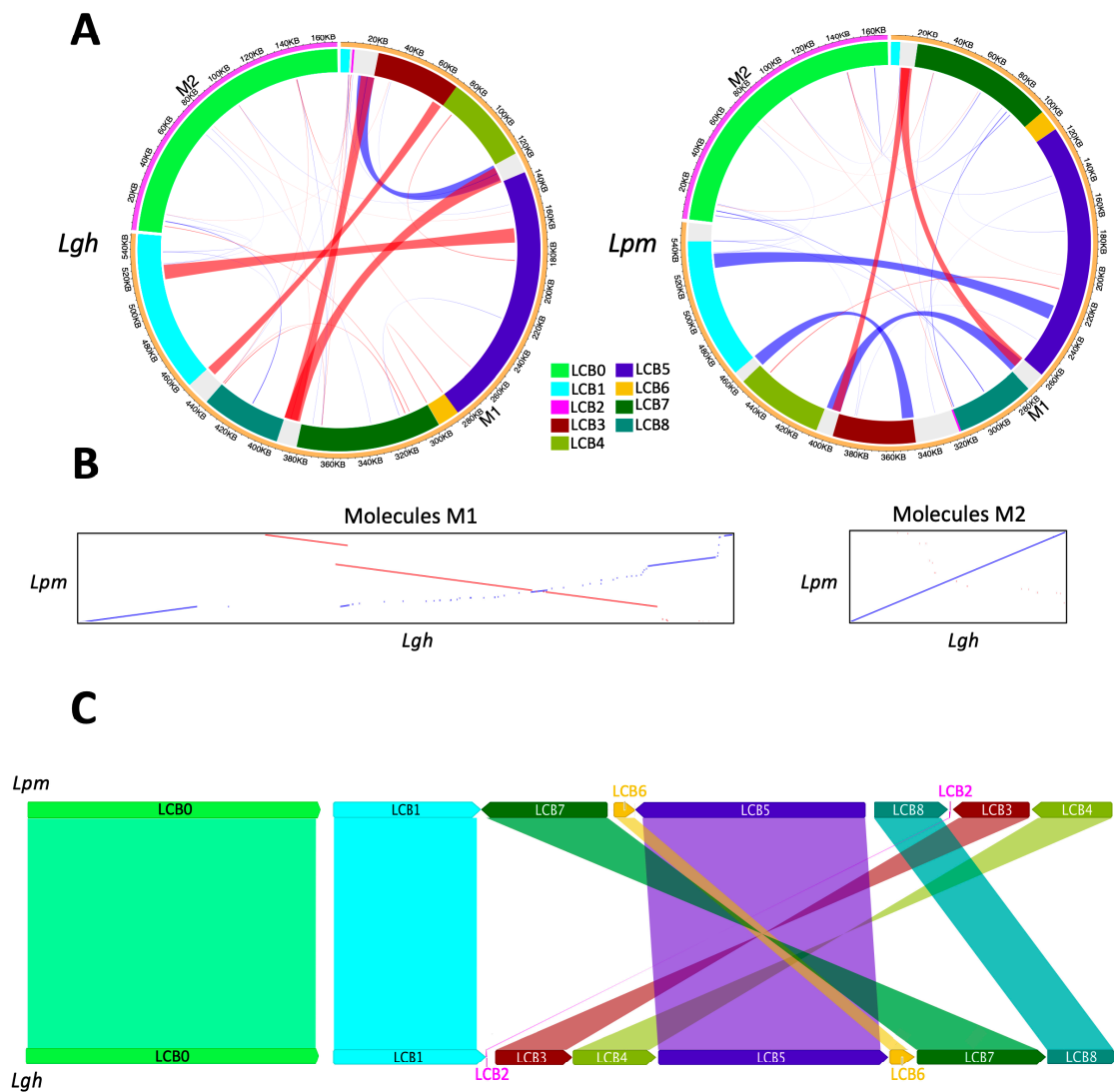


Figure 2. Structural comparisons between the *Ludwigia grandiflora* subsp. *hexapetala* (*Lgh*) and *Ludwigia peploides* subsp. *montevidensis* (*Lpm*) mitogenomes. (A) Repeats spanning at least 100 bp are illustrated with curved lines, with the width of the lines proportional to the size of the repeats. Inverted repeats are denoted in red, while direct repeats are indicated in blue. LCBs (locally collinear blocks) represent conserved regions between the *Lgh* and *Lpm* mitogenomes, as depicted in panel (C). (B) Comparisons of the M1 and M2 sequences of the two *Ludwigia* species were performed using LastZ: blue lines indicate common sequences in direct orientation, while red lines denote sequences that are common but in inverted orientation. (C) Locally collinear blocks (LCBs) (>1 kb) identified by ProgressiveMauve between the mitogenomes of *Lgh* and *Lpm*.

2.2.2. Protein Coding Gene Content

The mitogenomes of both *Ludwigia* species encompassed a total of 42–46 protein-coding genes, all located inside the LCBs, 68–70% of which were on the M1 molecule. Among these genes, 23 were involved in the oxidative phosphorylation (OXPHOS) machinery (Table 1), and of these, nine were associated with complex I (NADH dehydrogenase; *nad1–7*, *nad9*, and *nad4L*), with three trans-splicing genes (*nad1*, *nad2* and *nad5*) exhibiting distinct exon distributions. Exons of *nad2* were all localized on molecule M2, whereas exons of *nad5* were encoded on molecule M1, spread across LCB1 and LCB5, and exons of *nad1* were split between the two molecules. Specifically, exons 1, 4, and 5 resided in M1 (LCB1 and LCB8), while exons 2 and 3 were located in M2. Significantly, the *matR* gene was encoded within exon 1 of *nad1*, a phenomenon consistently observed across all angiosperm mitogenomes (PMID: 24600456). The genes of the other complexes were as follows: *sdh4* for complex II (succinate dehydrogenase), *cob* for complex III (cytochrome c reductase; three genes in complex IV (cytochrome c oxidase; *cox1–3*), and five in complex V (ATP synthase; *atp1*, 4, 6, 8, and 9). Additionally, there were four subunits dedicated to the biogenesis of cytochrome c (*ccmB*, *ccmC*, *ccmFN*, and *ccmFC*), along with genes encoding maturase (*matR*) and transport membrane protein (*mttB*). These genes were distributed across the different LCBs. Furthermore, 14 non-core genes encoding small and large subunits of ribosome proteins were found to be present (Table 1). We identified seven cis-splicing protein-coding genes (*ccmFc*, *cox2*, *nad4*, *nad7*, *rpl2*, *rps3*, and *rps10*) and annotated seven protein-coding genes that originated from their respective plastids. The genes *petN* (encoding cytochrome b6f), *psbZ* (involved in photosystem II), and *rps14* (ribosomal protein S14) were transferred intact in both the *Lgh* and *Lpm* mitogenomes, whereas *ndhJ* and *ndhC* (NADH dehydrogenase), *atpF* (ATP synthase), and *rps2* (ribosomal protein S2) were intact in *Lpm* mtDNA but underwent pseudogenization in the *Lgh* mitogenome (Table S3).

Table 1. Mitochondrial protein coding genes annotated in the *Ludwigia grandiflora* subsp. *hexapetala* (*Lgh*) and *Ludwigia peploides* subsp. *montevicensis* (*Lpm*) mitogenomes. Numbers within parentheses indicate the number of exons for spliced genes. Trans-spliced genes are marked by “*”. Locally collinear blocks (LCBs) correspond to conserved regions between *Lgh* and *Lpm* mitogenomes, identified by ProgressiveMauve. The molecular weight and isoelectric point of protein were calculated using the Sequence Manipulation Suite. Hypothetical localization was estimated by Plant-mPLoc.

Gene	LCB	Molecule	Size (in bp)	Molecular Weight (in kDa)	Protein Isoelectric Point	Hypothetical Localization
Complex I (NADH dehydrogenase)						
<i>nad1</i> (5*)	LCB0	M2	1008	36.81	10.10	Mitochondrion
	LCB1	M1				
	LCB8	M1				
<i>nad2</i> (5*)	LCB0	M2	1518	55.34	9.28	Mitochondrion
<i>nad3</i>	LCB3	M1	357	13.39	4.33	Mitochondrion
<i>nad4</i> (4)	LCB5	M1	1473	54.27	8.97	Mitochondrion
<i>nad4L</i>	LCB5	M1	303	10.94	6.70	Mitochondrion
<i>nad5</i> (5*)	LCB1	M1	1794	74.55	7.90	Mitochondrion
	LCB5	M1				
<i>nad6</i>	LCB1	M1	618	23.45	10.53	Mitochondrion
<i>nad7</i> (5)	LCB8	M1	1182	44.32	7.87	Mitochondrion
<i>nad9</i>	LCB5	M1	573	22.63	8.20	Mitochondrion
Complex II (succinate dehydrogenase)						
<i>sdh4</i>	LCB7	M1	393	15.25	10.76	Mitochondrion

Table 1. Cont.

Gene	LCB	Molecule	Size (in bp)	Molecular Weight (in kDa)	Protein Isoelectric Point	Hypothetical Localization
Complex III (cytochrome c reductase)						
<i>cob</i>	LCB5	M1	1182	44.04	7.52	Cell membrane
Complex IV (cytochrome c oxidase)						
<i>cox1</i>	LCB0	M2	1623	58.75	9.25	Mitochondrion
<i>cox2</i> (2)	LCB7	M1	780	29.09	5.53	Mitochondrion
<i>cox3</i>	LCB7	M1	798	29.54	7.53	Mitochondrion
Complex V (ATP synthase)						
<i>atp1</i>	LCB3	M1	1533	55.51	6.21	Mitochondrion
<i>atp4</i>	LCB0	M2	573	21.25	10.26	Mitochondrion
<i>atp6</i>	LCB8	M1	1044	38.30	4.80	Mitochondrion
<i>atp8</i>	LCB7	M1	480	18.19	9.50	Mitochondrion
<i>atp9</i>	LCB0	M2	330	11.61	9.09	Mitochondrion
Cytochrome c biogenesis						
<i>ccmB</i>	LCB4	M1	621	23.07	8.62	Mitochondrion
<i>ccmC</i>	LCB6	M1	723	26.92	10.98	Mitochondrion
<i>ccmFc</i> (2)	LCB4	M1	1500	56.40	10.74	Mitochondrion
<i>ccmFn</i>	LCB7	M1	1983	74.21	10.45	Mitochondrion
Maturase						
<i>matR</i>	LCB8	M1	1983	73.76	10.71	Chloroplast
Protein translocase						
<i>mttb</i>	LCB1	M1	768	29.07	7.75	Mitochondrion
Ribosomal large subunit						
<i>rpl2</i> (3)	LCB0	M2	1389	50.69	11.28	Mitochondrion
<i>rpl5</i>	LCB0	M2	564	21.55	5.89	Mitochondrion
<i>rpl10</i>	LCB4	M1	486	18.53	7.45	Cell membrane, chloroplast, and nucleus
<i>rpl16</i>	LCB0	M2	435	15.98	11.81	Mitochondrion
Ribosomal small subunit						
<i>rps1</i>	LCB0	M2	693	25.87	11.89	Mitochondrion
<i>rps3</i> (2)	LCB0	M2	1659	63.57	10.96	Mitochondrion
<i>rps4</i>	LCB0	M2	1047	41.18	11.63	Mitochondrion
<i>rps7</i>	LCB5	M1	447	17.04	11.21	Chloroplast and mitochondrion
<i>rps10</i> (2)	LCB7	M1	408	16.40	11.25	Mitochondrion
<i>rps12</i>	LCB3	M1	378	14.30	11.80	Chloroplast and mitochondrion
<i>rps13</i>	LCB0	M2	351	13.27	11.21	Mitochondrion
<i>rps14</i>	LCB0	M2	303	11.96	12.04	Mitochondrion
<i>rps19</i>	LCB0 LCB7	M2 M1	285	11.20	12.32	Chloroplast

2.2.3. RNA Editing

In the mitogenomes of *Lgh* and *Lpm*, 738 and 788 post-transcriptional C to U editing events, with a probability of more than 50%, were identified within the coding DNA sequences (CDS) (Table S4; Figures S1 and S2). Among these, the *psab*, *nad4*, *ccmB*, and *mttB* genes exhibited the highest numbers of RNA editing sites, with 49, 46, 45, and 41 sites, respectively. Approximately 80% of the post-transcriptional modifications were non-synonymous (589–623 non-synonymous modifications compared to 149–165 synonymous modifications in *Lgh* and *Lpm*, respectively). Eight and twelve RNA editing events led to the creation of stop codons in the mitogenomes of *Lgh* and *Lpm*. These stop codon occurrences were predominantly observed in the chloroplast genes (*atpE*, *atpI*, and *psaB* in both mitogenomes, and *ndhK* in the *Lpm* mitogenome only), resulting in the annotation of these genes as pseudogenes. Additionally, two RNA editing sites generated stop codons in two mitochondrial genes (*atp9* and *ccmFc*), resulting in their truncation.

2.2.4. Content of RNA Genes

In both *Ludwigia* mitogenomes, we identified and annotated a comprehensive set of 34 tRNA genes (Tables S3 and S5). This included 19 native mitochondrial tRNAs and an additional 15 tRNAs derived from the plastome. Among these, sixteen tRNA genes were unique whereas *trnM-CAU* was found in nine copies, *trnG-GCC* in three copies, and *trnN-GUU*, *trnP-UGG*, and *trnS-UGA* in two copies each. Three of them were cis-splicing tRNAs (*trnA-UGC*, *trnI-GAU*, and *trnV-UAV*), all of which originated from the plastome. As shown in Table S5, all tRNAs were located in LCBs, except for the three *trnM-CAU* copies present in the previously described 5253 bp repeated region. As described previously, our analysis revealed the presence of multiple copies of mitochondrial ribosomal RNA genes (*rrn5*, *rrn18*, and *rrn26*; Table S5) along with a singular copy of the plastid ribosomal RNA genes (*rrn4.5*, *rrn5*, and *rrn16*) (Table S3).

2.2.5. Chloroplast-Derived Fragments

In addition to the previously mentioned three to seven protein-coding genes, fifteen tRNA genes, and three rRNA genes, our analysis revealed twenty-four homologous fragments between the *Lgh* and *Lpm* chloroplast genomes (Figure 3; Table S3). Cumulatively, these fragments had a size of approximately 55 kb, contributing to about 7.8% of the entire mitochondrial genome length in both *Ludwigia* species. Notably, these insertions of chloroplast DNA were dispersed throughout the entire M1 and M2 molecules (Figure 3), indicating potential independent transfers or regions modified through subsequent recombination. The largest of these fragments (8051 bp), located on the M2 molecule, was 98.4% identical to the original chloroplastic region present in both the *Lgh* and *Lpm* plastomes. This included the three pseudogenes *atpA*, *atpH*, and *atpI*, which harbors a stop codon due to post-transcriptional RNA editing, and the two genes *atpF* and *rps2*, which were complete in *Lpm* and pseudogenized in *Lgh* as well as a segment of the *rpoC2* gene. All pseudogenizations were associated with point mutations that led to frameshifts. The next largest fragment, spanning 7.6 kb and located on the M1 molecule, exhibited a lower identity to the plastome (93%) and featured an intact *trnM-CAU* gene at its extremity, accompanied by degraded forms of the *ycf2* and *ycf15* genes. In addition, we identified a 6.1 kb region with 99.6% identity to cpDNA, encompassing three tRNA genes, the *rrn16* gene, and a fragment of *rrn23*.

Another region, 5.4 kb in length, exhibited more degradation in *Lgh* compared to *Lpm* (89.3% and 91.9% identity with cpDNA, respectively), resulting in the pseudogenization of the *ndhJ* and *ndhC* genes in *Lgh* only, whereas *ndhK* and *atpE* were pseudogenized in both mitogenomes. Finally, the ultimate two extended regions of plastid origin, measuring 4.8 and 3.6 kb, respectively, were identical in both species and corresponded to the original cpDNA copy (Table S3), albeit with gene fragments (*psaA*, *ycf1*, and *rrn23*) at their termini. The remaining plastidic-originated fragments, ranging in size from 1215 to 147 bp contained either complete tRNA sequences or, for the majority, fragments of protein-coding genes.

Some of these corresponded to remains of *rps12*, *atpB*, or *psaA*, present in two copies (Table S3).

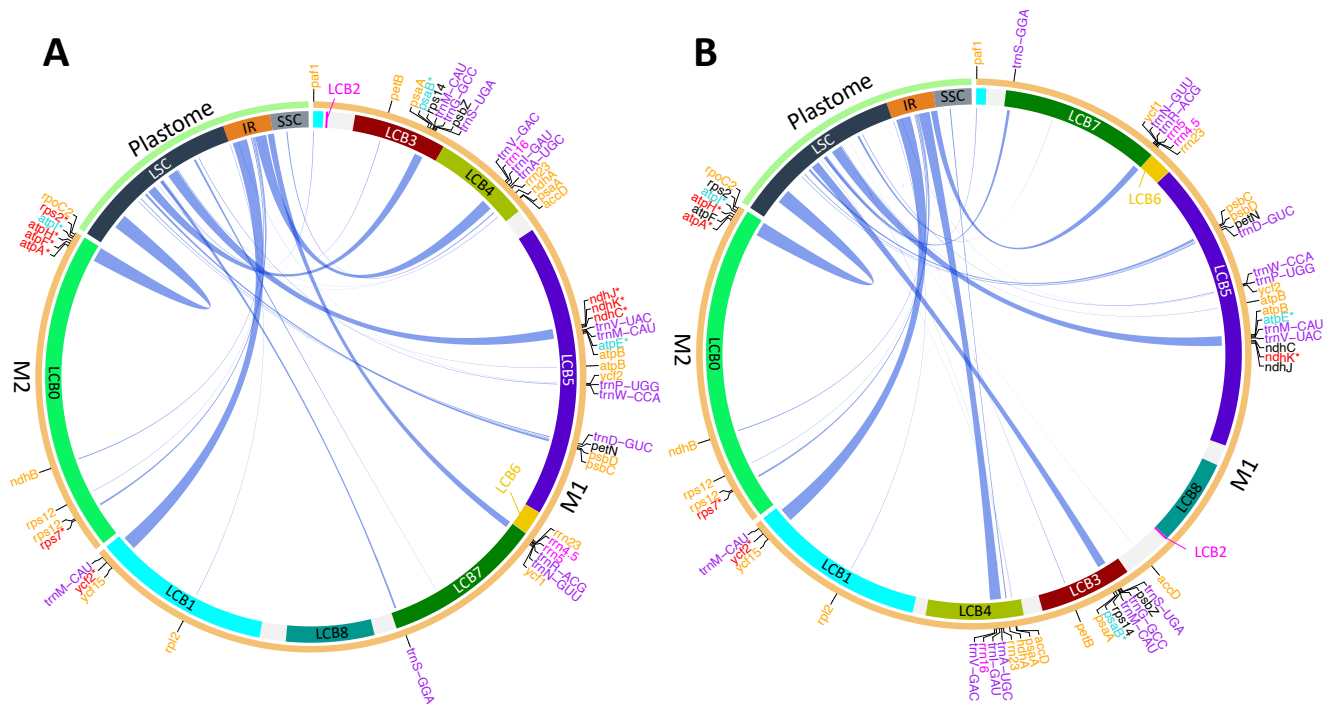


Figure 3. Maps illustrating the chloroplast sequence insertions (>80 bp) in the mitogenomes *Ludwigia grandiflora* subsp. *hexapetala* (*Lgh*; (A)) and *Ludwigia peploides* subsp. *montevidensis* (*Lpm*; (B)). Curved lines connect chloroplast insertions in the mitogenome to their respective origins in the plastome and the line width is proportional to the insertion size. Locally collinear blocks (LCBs) represent conserved regions between the *Lgh* and *Lpm* mitogenomes, identified by ProgressiveMauve. Plastome regions including large single copy (LSC), inverted repeat (IR; represented once for simplification), and small single copy (SSC), are indicated. Genes that originated from the plastome are color-coded: pseudogenes in red, pseudogenes due to RNA editing in blue, fragment genes in orange, complete protein coding genes in black, tRNA genes in purple, and rRNA genes in magenta. * Correspond to pseudogene.

2.2.6. Fragments of Mobile Genetic Elements

Four main types of endogenous fragments of mobile genetic elements were widespread in both the M1 and M2 *Ludwigia* mitogenomes: mitovirus NERVES (nonretroviral endogenous RNA virus elements), all three classes of nuclear retrotransposons, the *Tyl/copia*, *Ty3/gypsy*-, and non-LTR/LINE-families, and plasmid derivatives (Table S6). For mitovirus NERVES, all fragments corresponded to partial or pseudogenized copies of the RNA dependent RNA polymerase (RdRP) of mitoviruses and were identical in *Lgh* and *Lpm*. The retrotransposon fragments integrated into the *Ludwigia* mitochondrial genomes were dispersed throughout the intergenic regions of the mitogenomes without showing a significant preference for integration hotspots. No complete copy was observed for any of the mobile elements identified.

2.2.7. Repeated Sequences and Recombination

Three non-tandem repeats exceeding 1 kb in size were detected within both the *Ludwigia* M1 mitogenomes, whereas none were detected in the M2 molecules (Figure 4). The first repeat encompassed a central region spanning 5641 bp (called “Core R1”; Figure 4A), was found in triplicate within both *Ludwigia* M1 molecules, and carried a gene for both 26S rRNA (*rrn26*) and the initiator methionine tRNA *trnM-CAU*. This core region exhibited a 99.9% identity between *Lgh* and *Lpm*, except for a divergent trinucleotide at position 386 (TCC in *Lgh*, GGA in *Lpm*) and a dinucleotide at position 5261 (TC in *Lgh*, GA in *Lpm*). In the

vicinity of this central “Core R1” area, we identified three scenarios: “R1-right” (Figure 4A), corresponding to an expansion in the 3′ region, which included both the “Extension” and “Right” regions present in *Lpm* and *Lgh* as well as an additional region called “Right_lgh” exclusive to *Lgh*; “R1-left” (Figure 4A), corresponding to an additional 5′ region (called “Left” for both and a small region unique to *Lgh* called “Left_lgh”), which was added to the “Core R1” part in *Lpm* and to the “Core_R1” and “Extension” regions in *Lgh*; and finally, “R1-left-right” (Figure 4A), which combined both the “R1-right” and “R1-left” regions. The second and third repeats (R2 and R3) were classic repetitive sequences, each appearing twice in the *Lpm* and *Lgh* M1 molecules. The R2 repeats, which harbored a solitary gene encoding *trnG-GCC* at their 3′ termini, spanned 9384 and 9376 base pairs in *Lgh* and *Lpm*, respectively. These repeats showcased a remarkable 99.9% identity across the two *Ludwigia* mitogenomes. Notably, copies found in *Lgh* possessed an additional eight bases and exhibited two single nucleotide polymorphisms (SNPs) compared to those present in *Lpm*. R2 repeats were in direct orientation in *Lpm* but were reversed in *Lgh*. *Lpm* repeats R3 were 1554 nt longer than *Lgh* repeats R3, but their common regions of 7817 nt were highly conserved (99.9% identity, 5 SNPs). Like R2, the R3 repeats were in direct orientation in *Lpm* and reversed in *Lgh*. Placing these repeats within the framework of the locally collinear blocks (LCBs) shared between *Lpm* and *Lgh*, it was apparent that all R1 repeats and one of the R3 repeats were positioned in LCBs within non-conserved regions, as depicted in Figure 4B. Meanwhile, the second iteration of R3 resided within LCB3, close to a fragment of chloroplastic origin. Additionally, the two R2 repeats were identified on LCB1 and LCB5.

By using these large repeated sequences, multiple alternative conformations were tested. In the *Lgh* mitogenome, an alternative form of the M1 molecule was found with a probability of 50%, mediated by R1-right (Figures 4B and 5). In this case, the M1 molecule was divided into two smaller molecules of 435 and 109 kb. Some other conformations mediated by R2, R1-core (The conserved part between R1-left and R1-right), and R3 repeats could not be completely validated but might also exist (Figures 4B and 5). Recombinations mediated by these three repeats led to all molecules of the same size as the original M1 mitogenome, each with a specific inversion (Figure 5). By combining all of these recombinations, three molecules of 109, 174, and 261 kb could be obtained. In the *Lpm* mitogenome, an alternative form of the M1 molecule was identified as mediated by R1-right like in *Lgh* (Figure 6). This form was composed of two molecules of 425 and 130 kb. Due to the lack of reads, we could not identify as many putative forms as in the *Lgh* M1 molecule, and only one was found with R1-core (Figure 6). This consisted of an M1 molecule of 555 kb with an inversion. A combination of the two rearrangements may also exist.

A total of 266–294 non-tandem repeats (Table S7; Figure S3) of at least 30 bp were identified in the *Ludwigia* mitogenomes and corresponded to 127–169 direct repeats and 125–139 reverse repeats. Non-tandem repeats represented 7–7.4% (50,752–52,800 bp) of each mitogenome. Most of them (161–185; 60.5–62.9%) were only found on M1, while 34–36.8% (98–100) of repeats were found between the M1 and M2 molecules and 2.6–3% (7–9) were only on the M2 molecule. The majority of repeats (192–220; 72.2–74.8%) were between 30 and 59 bp (Figure S3). Only 10.2–13.5% (30–36) of repeats were larger than 100 bp. A total of 25 and 22 tandem repeats were found in the *Lgh* and *Lpm* mitogenomes (20–24 on M1 and one or two on M2; Table S8). These repeats were in non-coding regions, except for the three of 18-period size present in each of the three copies of the *rrn26* genes.

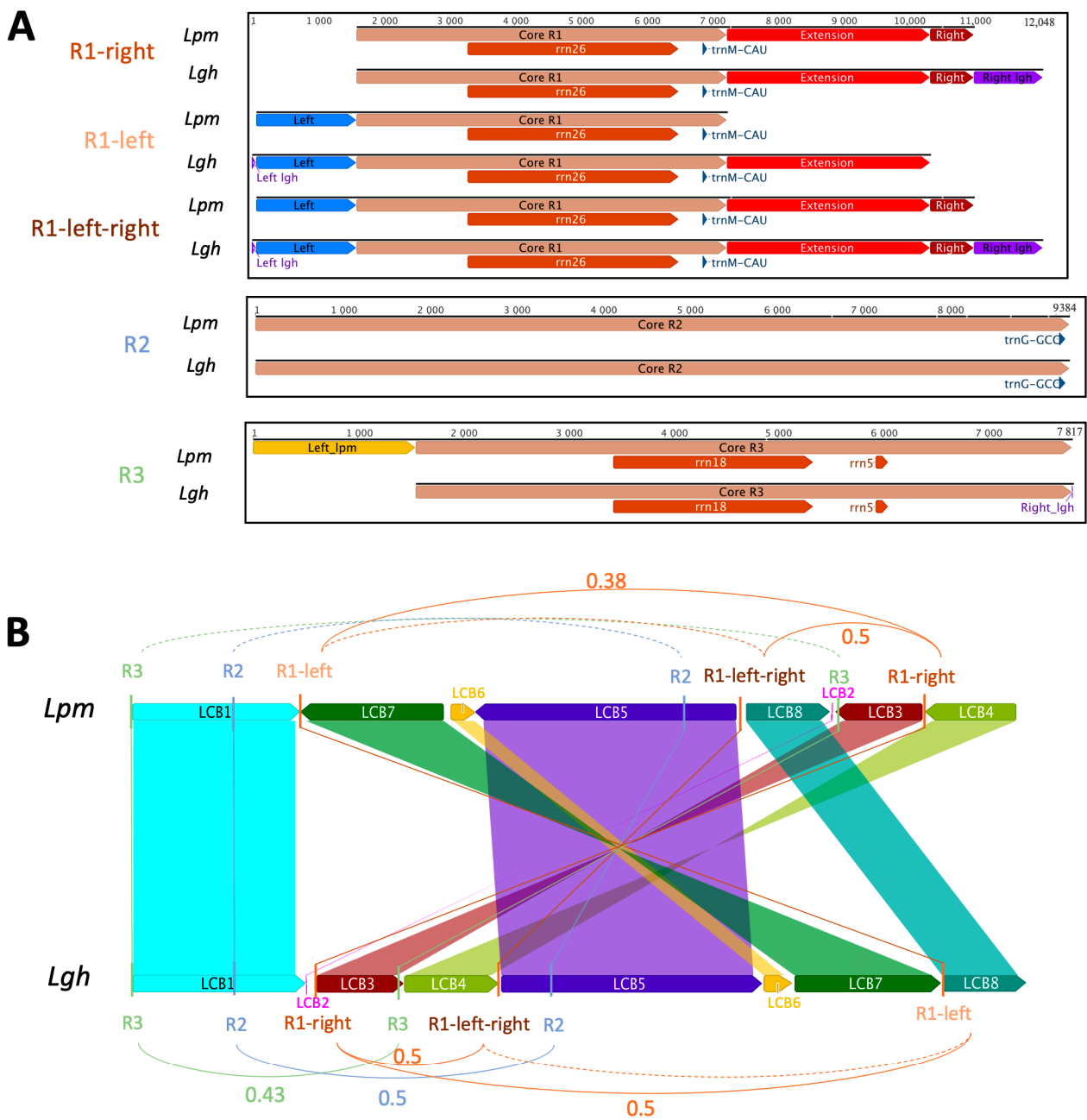


Figure 4. This representation of repeats exceeding 1 kb in the mitogenomes of *Ludwigia grandiflora* subsp. *hexapetala* (*Lgh*) and *Ludwigia peploides* subsp. *montevidensis* (*Lpm*). “R” denotes repeat. R1-left-right combines R1-right and R1-left. (A) The five repeats exceeding 1 kb in the mitogenomes of *Lpm* and *Lgh*. Core sequences represent conserved sequences among repeats. “Extension” sequences were present in all repeats except R1-left in *Lpm*. Left and right sequences were unique to R1-left and R1-right, respectively, and both were present in R1-left-right. Purple and yellow sequences were unique to *Lgh* and *Lpm*, respectively. (B) Sequence comparison of M1 molecules between the two *Ludwigia* species and their repeats. Locally collinear blocks (LCBs) represent conserved regions between the *Lgh* and *Lpm* mitogenomes, identified by ProgressiveMauve. Recombination frequencies are specified when >0, while a dotted line indicates the absence of recombination.

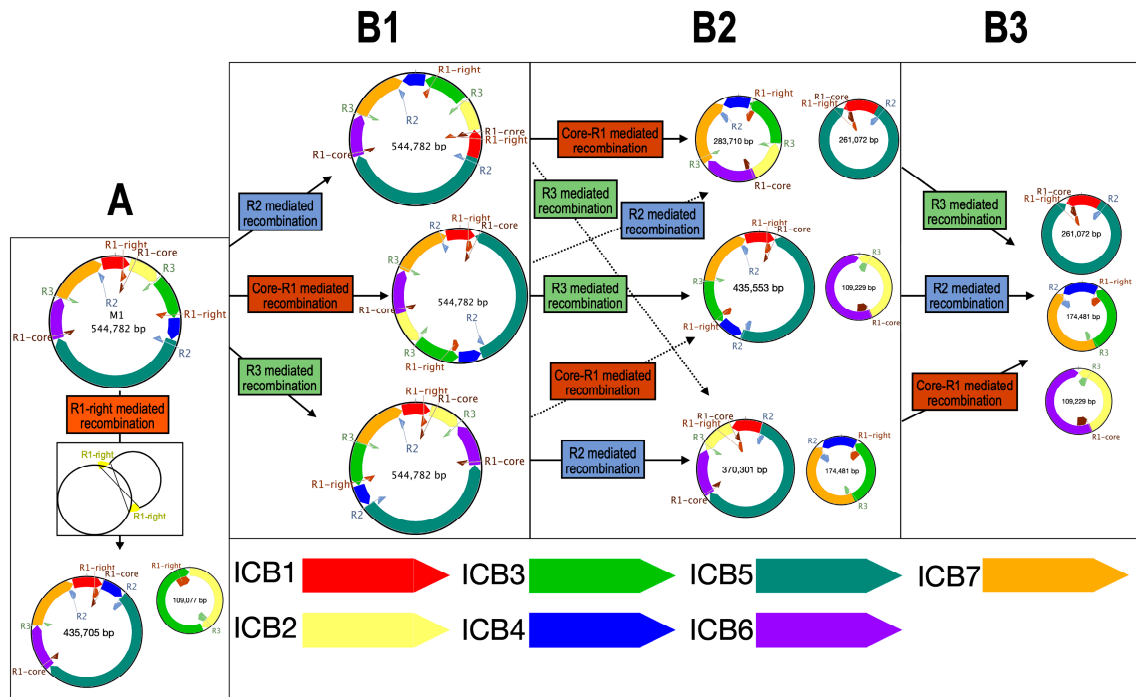


Figure 5. Schematic of mitogenome *Lgh* recombinations: (A) Validated conformations after M1 molecule rearrangement in *Ludwigia grandiflora* subsp. *hexapetala* mitogenome. (B1–B3) Different steps of the feasible successive recombinations. Only intramolecular rearrangements are depicted. “R” denotes repeat. ICB (intramolecular collinear block) represents regions conserved among all forms.

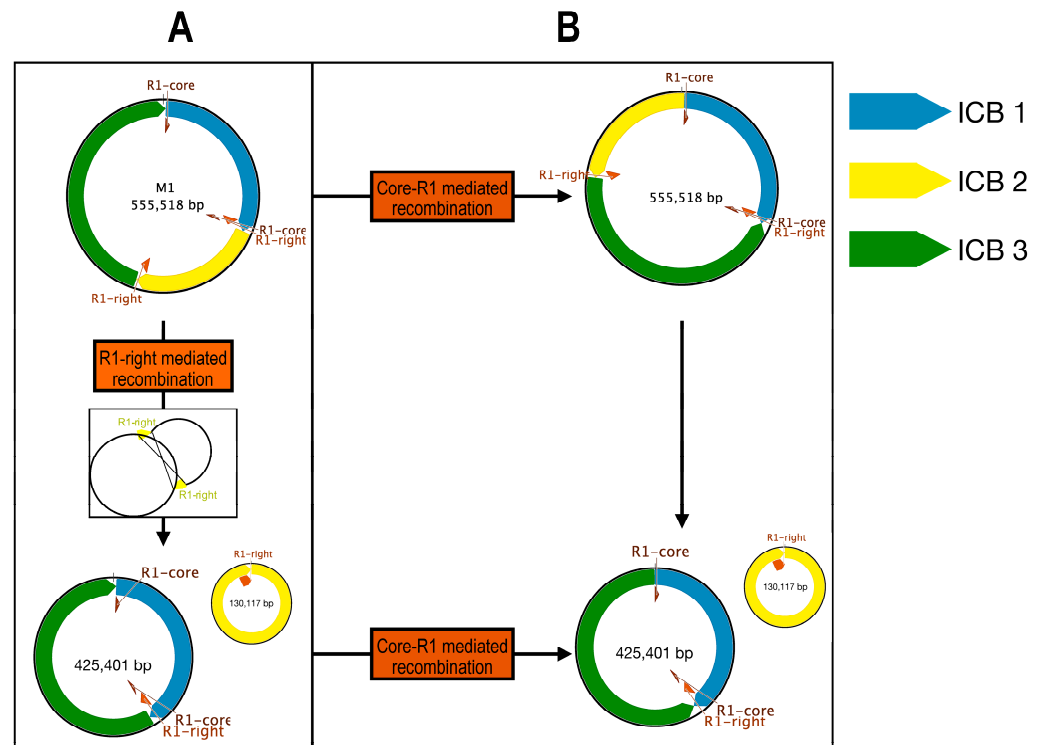


Figure 6. Schematic of mitogenome *Lpm* recombinations: (A) Validated conformations after M1 molecule rearrangements of *Ludwigia peploides* subsp. *montevidensis* mitogenome. (B) Putative conformations after M1 molecule rearrangements of *Ludwigia peploides* subsp. *montevidensis* mitogenome. Only intramolecular rearrangements are depicted. “R” denotes repeat. ICB (intramolecular collinear block) represents regions conserved among all forms.

In the *Ludwigia* mitogenomes, 183–155 SSRs were identified (137–117 and 46–38 found on the M1 and M2 molecules, respectively; Figure S4). The majority of SSRs were monomeric (66–44) and tetrameric (63–56). Monomeric repeats were almost entirely composed of A or T (30–20 A and 34–24 T). A total of 32–36 dimeric, 15–13 trimeric, and 7–6 pentameric SSRs were identified. Dimeric repeats of TA were the third most common SSR (9–10) after monomeric A and T repeats.

2.3. Gene Comparisons and Phylogeny

Relative synonymous codon usage (RSCU) was calculated using MEGA in the *Lgh* and *Lpm* mitogenome CDS on 11,595 and 12,141 codons, respectively (Figure 7; Table S9). A total of 30–29 codons had an RSCU > 1, which means that they were used more often than another synonymous codon. Among them, the most preferentially used codons were A-ended or U-ended, except for the UGG leucine codon, showing a strong bias in the third letter of the codon. In both mitogenomes, the UGA stop codon had the biggest RSCU value (1.6–1.64). In *Lgh*, the smallest RSCU value was for the UAG stop codon (0.29) while in *Lpm*, it was the GCG alanine codon (0.42).

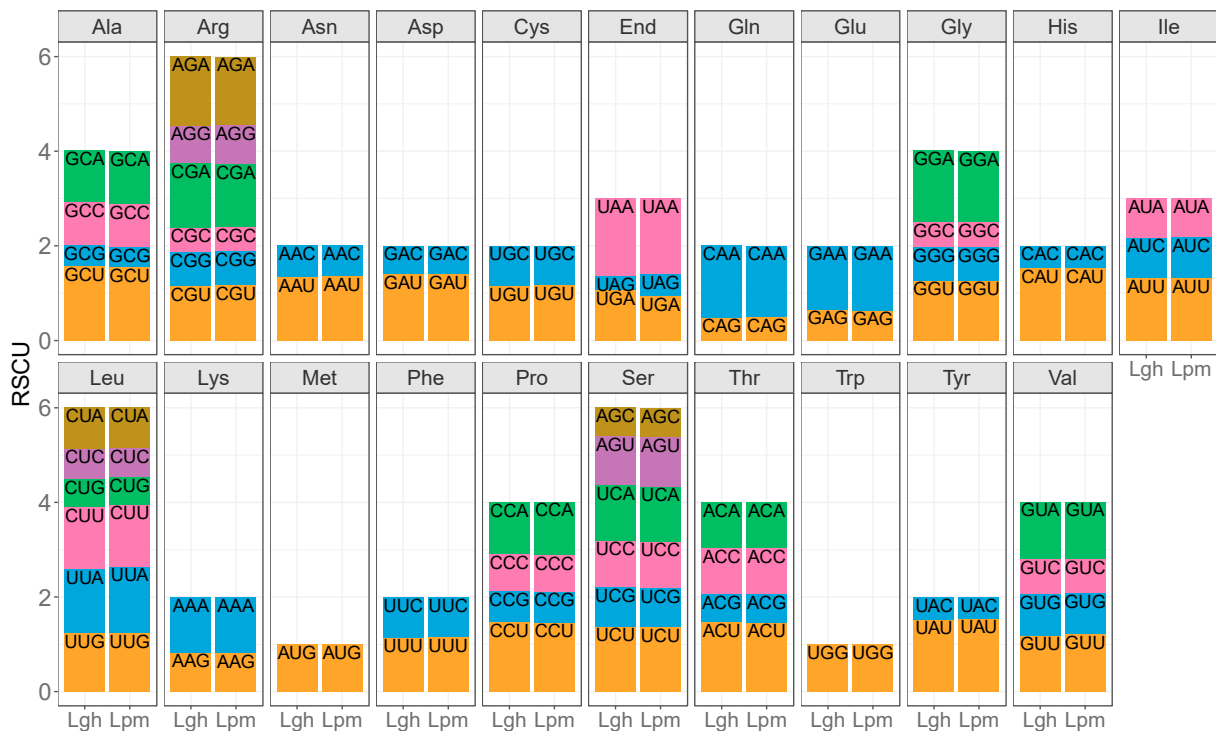


Figure 7. The relative synonymous codon usage (RSCU) bias in *Ludwigia grandiflora* subsp. *hexapetala* (*Lgh*) and *Ludwigia peploides* subsp. *montevidensis* (*Lpm*) mitogenome CDS (sequences of protein coding genes).

To identify genes under selection, K_a (non-synonymous substitution rate) and K_s (synonymous substitution rate) were calculated on 31 genes by comparing *Ludwigia* mitogenomes (these genes were identical in *Lgh* and *Lpm* mitogenomes) to each of the other 14 assembled Myrtales mitogenomes one by one, and to the *Geranium maderense* mitogenome as an outgroup. Genes under positive selection (not conserved) will have a K_a/K_s value above one, whereas genes under negative selection (conserved) will have a K_a/K_s value under one. A K_a/K_s ratio value of one points to a neutral selection. The *rps1* gene is the only gene showing a $K_a/K_s > 1$ in the majority of species studied (Figure 8, Table S10). Moreover, the *ccmB* and *mttB* genes were also under negative selection in *Lgh* and *Lpm* in seven and five species, respectively. However, the majority of K_a/K_s ratios (92%) were under 1, showing the conservation of most genes.

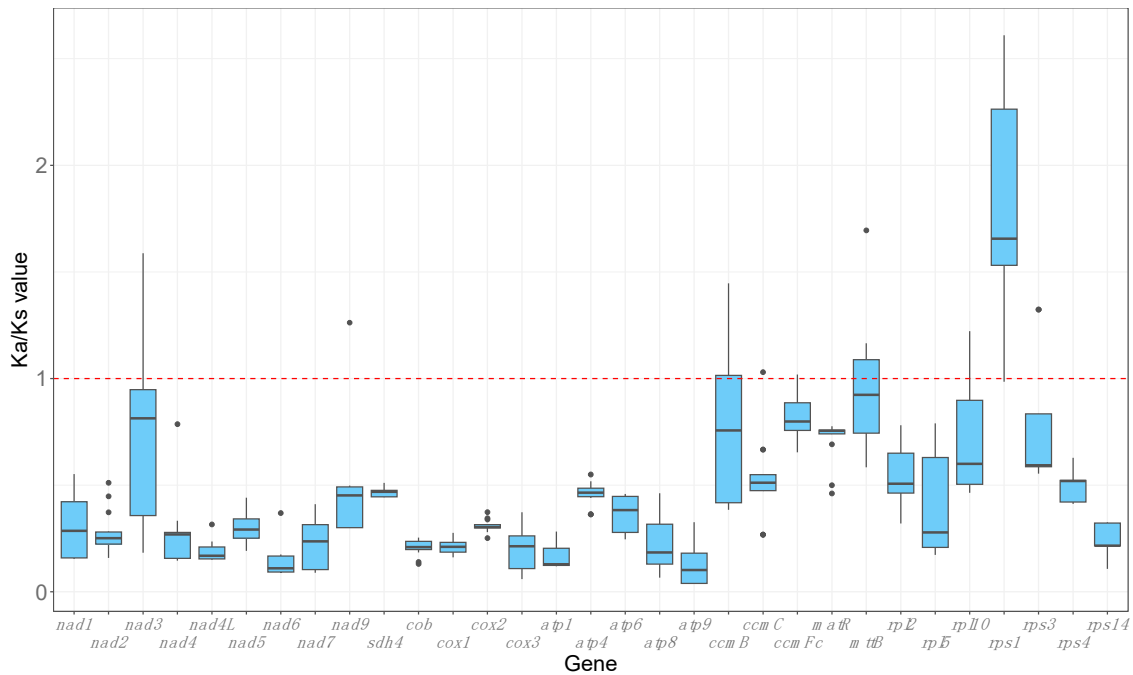


Figure 8. Graphical representation of pairwise comparisons of non-synonymous substitution rate (Ka) and synonymous substitution rate (Ks) ratios among species from the orders Myrtales and Geraniales, utilizing *Ludwigia grandiflora* subsp. *hexapetala* and *Ludwigia peploides* subsp. *montevidensis* as references, across 31 mitochondrial genes. Genes were included if they were present in at least 14 out of the 16 selected species.

Sixteen conserved protein-coding genes were concatenated and aligned to make a phylogenetic tree of the Myrtales order with *Geranium maderense* used as an outgroup (Figure 9). The tree fit with the phylogeny in the Myrtales order. The outgroup *G. maderense* was separated from the Myrtales order species. Species from the same family were grouped. The four Onagroideae subfamily species were clustered together within the Onagraceae family. Finally, the two *Ludwigia* species were on the same branch.

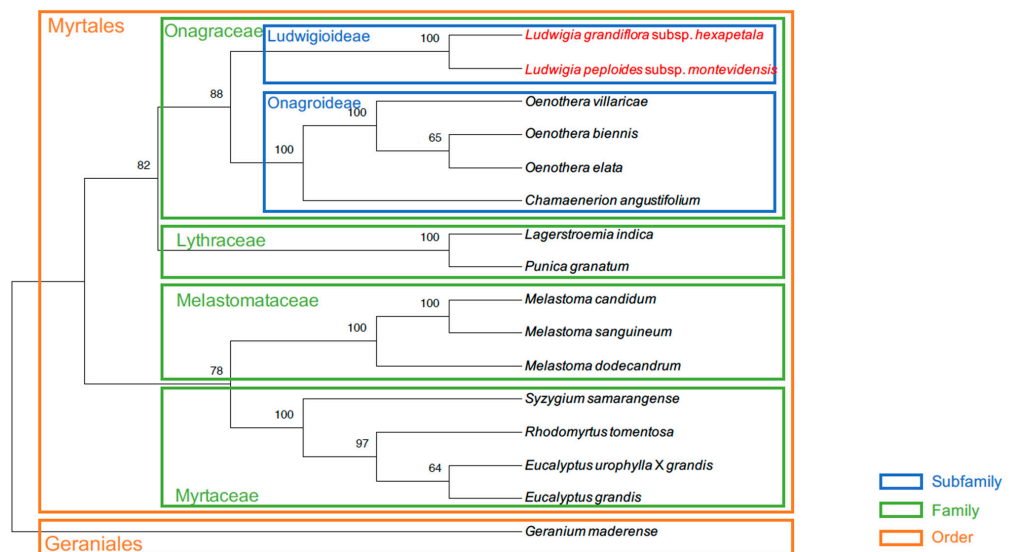


Figure 9. Phylogenetic tree of the Myrtales order constructed using 16 protein-coding genes via the maximum likelihood method with a bootstrap value of 1000. *Geranium maderense* was designated as the outgroup. Branch frequencies are denoted by numbers on branches. The *Ludwigia* species sequenced in this paper are highlighted in red.

Gene compositions were compared in the 15 Myrtales order species and *G. maderense* mitogenomes (Figure 10). Sixteen genes were common between all of the studied mitogenomes. Ten genes were only missing in one mitogenome. Among them, nine were missing in *R. tomentosa* and one in *E. urophylla* × *grandis*. Mitochondrial genes missing in more than one mitogenome were ribosomal subunit genes (*rpl16*, *rps7*, *rps10*, *rps12*, *rps13*, and *rps19*), *ccmFn* and *sdh3*. Out of all species, 83% of the identified chloroplast protein coding genes were incomplete. Only five species contained one to nine complete chloroplast genes: *Lgh*, *Lpm*, *M. candidum*, *M. sanguineum*, and *E. grandis*. These genes coded for different functions: PSI assembly factor (*ycf3*), photosystem II (*psbM* and *psbZ*), ATP synthase (*atpF*, *atpH*, and *atpI*), cytochrome b6f (*petN*), NADH dehydrogenase (*ndhB*, *ndhC*, and *ndhJ*), transcription (*rpoC1*), ribosomal large subunit (*rpl14*, *rpl22*, and *rpl23*), and ribosomal small subunit (*rps2* and *rps14*). Each of these genes were present in only one or two species, always of the same genus.

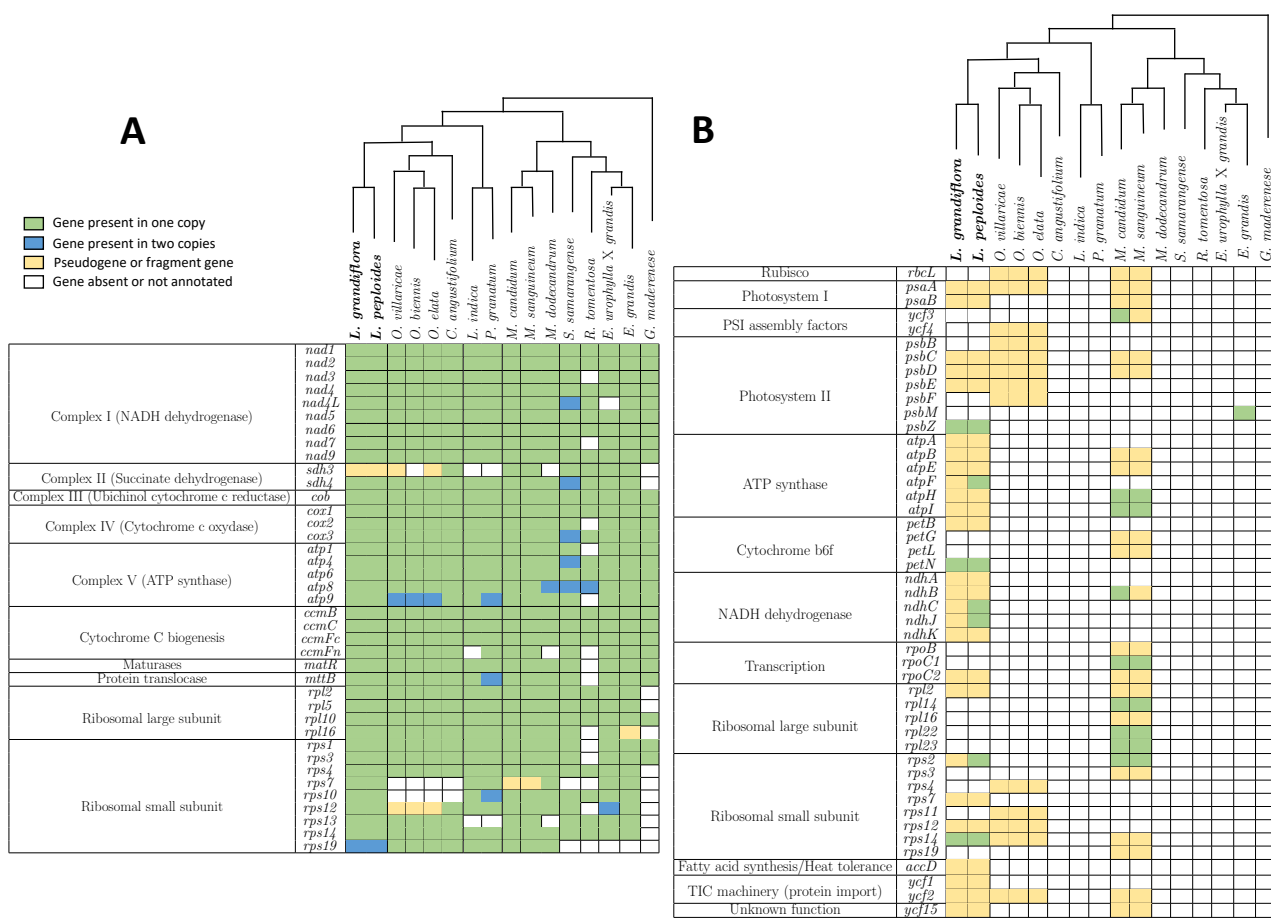


Figure 10. Gene compositions of mitogenomes from 15 Myrtales and one Geraniales order species. Panel (A) displays the mitochondrial genes, while Panel (B) shows the chloroplast genes. The phylogenetic trees presented were based on 16 protein-coding genes and constructed using the maximum likelihood method (refer to Figure 9). The *Ludwigia* species sequenced in this paper are highlighted in bold. Color boxes correspond to presence/absence panel (see legend in figure).

3. Discussion

In this study, *Lgh* and *Lpm* mitogenomes were assembled de novo from the SR or /and LR sequences. For *Lgh*, the best assembling combination was found to be SPAdes and Flye. Two circular molecules called M1 and M2 were obtained for both mitogenomes and had similar sizes (544,782 bp and 555,518 bp for M1 and 166,796 bp and 158,656 bp for *Lgh* and *Lpm*, respectively). For these two closely related species, their mtDNAs showed similarities in their structures and sequences for both the M1 and M2 molecules. While the

M2 molecules were highly similar (99.5%) and colinear, the M1 molecules from *Lpm* and *Lgh* mitogenomes contained eight conserved blocks including three inverted blocks.

3.1. Comparison of Plant Mitogenomes in Myrtales Order

Plant mitogenomes vary greatly in size, ranging from 66 kb to 12 Mb. In the Myrtales order, which comprises 9 families, 399 genera, and about 13,000 species [88], only the mitochondria genomes of 17 species have been sequenced. Most of these sequenced mitogenomes are around 400 to 450 kb in size, as follows for the Lythraceae family (*Punica granatum* (404,807 bp) [89], *Lagerstroemia indica* (333,948 bp; NC 035616), *Trapa incisa* (381,774 bp; NC 086691), *T. bicornis* (383,262 bp; NC 086690)); Myrtaceae family (*Rhodomyrtus tomentosa* (400,482 bp) [90], *Eucalyptus grandis* (478,813 bp) [91], *E. camaldulensis* (463,134 bp) [92], *E. urophylla* × *grandis* (481,982 bp; OQ947725-OQ947727), *Syzygium samarangense* (530,242 bp; NC 079700)); Melastomataceae family (*Medinilla magnifica* (377,864 bp) [93], *Melastoma candidum* (391,595 bp; NC 071383), *M. sanguineum* (395,542 bp; NC 071384), and *M. dodecandrum* (411,954 bp) [94]); and the Onagraceae family (*Oenothera villaricae*, *O. biennis* and *O. elata*, with respective lengths of 408,260, 424,132, and 418,451 bp [37]) and *Chamaenerion angustifolium* (495,176 bp; OX328283)). For *Lgh* and *Lpm*, however, the total length of mitogenomes was notably higher, standing at approximately 700 kb (711,578 bp and 722,518 bp, respectively), and even the largest M1 molecules alone exhibited a greater size than that of the mitogenomes of the previous 17 species.

The origin of this variation in size between different plant mitochondrial genomes is multifaceted, encompassing the proliferation of repeat elements, the integration of foreign DNA, and the acquisition or loss of large intragenic segments [95]. Homologous recombination has led to rearrangements that play a significant role in plant mitogenome evolution [96]. In their study, Zhou et al. (2023) conducted pairwise comparisons of the mitogenomes of five Myrtales species and observed sequence similarities ranging from 17% to 36% [94]. Notably, the highest percentage of shared sequences was observed between *M. magnifica* and *M. dodecandrum*, which belong to the same family. Similarly, the comparison of the mitogenomes of *E. camaldulensis* and *E. grandis* revealed a distinct alteration in repeat sequences, shedding light on evolutionary changes between the two genomes [92]. In the mitogenomes of *Lgh* and *Lpm*, the smaller M2 molecules exhibited colinearity and shared 99.5% identity, whereas the larger M1 molecules displayed rearrangements in five inter-LCB regions (size > 1 kb) between the two species, although no alterations were observed in the repeat sequences. One explanation for this result could be the genomic proximity of these closely related species, a hypothesis supported by the fact that *Lpm* is one of the maternal progenitor species of *Lgh* [33]. Given that most plants inherit mitochondrial genomes from the maternal parent, the minimal difference observed between the two mitogenomes might signify the evolutionary divergence between *Lpm* and *Lgh*.

3.2. Existence of Multi-Chromosomal Mitogenomes

Initially, most plant mitogenomes were depicted as single circular molecules. However, the increased availability of plant mitogenomes has unveiled a greater complexity, with multi-chromosomal mitogenomes identified in many species. However, research by Wu et al. (2020) indicates that multiple chromosome structures are present in only about 10% of sequenced plant mitogenomes [95]. Significantly, the mitochondrial genomes of *Lgh* and *Lpm* stand out as inaugural examples of multiple mitogenome structures documented within the Myrtales order. In this context, Bi et al. (2020) posed an intriguing question regarding the mechanisms that regulate and control the replication and segregation of multi-chromosomal mitogenomes during cell division [97]. In our study, we observed that for *Ludwigia*, sequence coverage was consistent for both the M1 and M2 molecules, indicating an equal number of copies of both molecules in the cell and therefore the likelihood of equi-replication.

Recently, Yang et al. (2022) reconstructed the mitogenome of *Salvia miltiorrhiza*, which was initially assembled as a single circular molecule through short-read sequence analy-

sis [98]. By employing both long-read and short-read technologies, the researchers identified two mitochondrial chromosomes, highlighting the value of combining both sequencing technologies. This finding reinforces our research, demonstrating that the integration of short-read (SR) and long-read (LR) sequences leads to improved mitogenome assembly. Notably, SR sequences, which are shorter than the repeat size, may not fully span the repeats independently, highlighting the necessity of combining both technologies for comprehensive assembly.

The presence of large repeat sequences (longer than 1 kb) has been shown to induce structural changes or isomerization in plant mitogenomes [58,99]. Hence, recent studies suggest the presence of one to two dominant forms of sub-genomic mitochondrial chromosomes, with repeat sequences capable of mediating recombination, leading to a diverse array of minor conformations. For instance, in *Gelsemium elegans*, four pairs of repeat sequences have been shown to facilitate the formation of one major and five minor conformations of the mitochondrial genome [100]. In *Salvia miltiorrhiza*, Yang et al. (2022) discovered two mitochondrial genomes and identified multiple minor conformations resulting from homologous recombination mediated by nine repeat sequences [98]. In our study, we identified two dominant forms, M1 and M2, and validated, through long-read (LR) sequence mapping, the existence of two minor conformations for the M1 molecule in both *Lgh* and *Lpm*, mediated by three repeat sequences. Other configurations are theoretically possible, but could not be confirmed with the sequences available.

3.3. Plastid DNA Transfer in *Lpm* and *Lgh* Mitogenomes

Inter-organelle DNA transfer can occur between the mitochondria and plastids, although it is extremely rare in the direction of mitogenomes to plastids. The plastome sequences represent 0% to 10% of the mitochondrial genome. The presence of chloroplast sequences and the number of genes in mitochondrial genome are very fluctuant. Recently, Wang et al. (2018) demonstrated that plastid-derived DNA fragments comprised 10.34% of the entire mitogenome length of *P. micranthus*, encompassing twelve intact protein-coding genes, three tRNAs, and twenty-nine pseudogenes [101]. Similarly, Tang et al. (2024) observed for *Acorus tatarinowii* and Yang et al. (2024) for *Haloxylon Ammodendron* that homologous sequences between the mitochondrial genome and the chloroplast genome accounted for 3.78% and 2.49%, respectively [102,103]. In our study, both the *Lgh* and *Lpm* M1 and M2 molecules contained plastome-origin sequences, encompassing 21–25 intact genes and constituting 7.8% of the mitochondrial genome length, which is quite a high percentage compared to the studies cited above, particularly given the large size of *Ludwigia* mitogenomes.

RNA editing is a post-transcriptional mechanism in higher plant organelles, and contributes to improved protein folding [104]. The conversion of C-to-U is the main form of RNA editing, as we observed in *Lgh* and *Lpm* [105]. RNA editing plays a crucial role in various plant developmental processes and could serve as a vital mechanism in adapting to a changing environment, allowing plants to acclimatize and survive [106]. Given the remarkable plasticity of both *Ludwigia* species to thrive in aquatic to dry environments [23], the role of RNA editing in facilitating this acclimatization warrants investigation. We identified the highest number of RNA editing sites among all mitochondrial genes in the *psab*, *nad4*, *ccmB*, and *mttB* genes. Similarly, Lu et al. (2024) reported elevated RNA editing sites in the *ccmB* and *nad4* genes in *Syzygium samarangense* [107]. Additionally, Xiong et al. (2017) investigated the responses of six mitochondrial genes including *ccmB* to oxidative stress, commonly associated with abiotic stresses like drought [108]. They demonstrated that differences in the RNA editing of *ccmB* could modulate plant performance responses under oxidative stress.

To summarize, the *Lpm* and *Lgh* mitogenomes consisted of two circular molecules with similar structures and sequences. The presence of large repeat sequences could result in multiple minor conformations. Plastome-origin sequences accounted for 7.8% of both mitochondrial genomes. A high number of RNA editing was observed.

4. Materials and Methods

4.1. Plant Material and Sequencing Data

To avoid genetic variability, *Lgh* and *Lpm* plants used for the sequencing were obtained using vegetative reproduction from a single fragment of stem with buds for each species. DNA extractions were carried out from new buds according to a specific protocol for extracting DNA long fragments [32]. DNA fragments were sequenced using Illumina MiSeq for short reads (SRs) and Oxford Nanopore GridION for long reads (LRs) at the Genome Transcriptome Facility of Bordeaux (INRAE-UMR 1202 BIOGECO, Bordeaux, France), as described in Barloy-Hubler et al. (2023). Both LR and SRs were available for *Lgh* but only LR for *Lpm*. Quality controls and filtering were performed using Guppy v4.0.14 for the nanopore reads and fastp v0.20.0 for the MiSeq reads [109]. Short reads were then paired and merged using BBMerge [110].

4.2. Assembly Strategies and Annotations

For the assembly of the *Lgh* mitogenome, SRs were corrected with LR using SPAdes and LR with SR using Ratatosk [111,112]. Subsequently, Flye was used to assemble the corrected LR, Megahit to assemble the corrected SR, and SPAdes for the corrected SR-LR hybrid assemblies [113,114].

To recruit mitochondrial reads, we developed a strategy centered on the coding DNA sequences (CDSs) due to their high conservation across species [57]. Therefore, all mitogenomes from the malvids clade were retrieved from the NCBI RefSeq database, and each CDS was identified, aligned, and a consensus sequence was constructed to serve as references to map corrected LR. LR selected in this manner were then assembled using the 'De Novo Assemble' tool within Geneious Prime 2023 (<http://www.geneious.com/>, accessed on 28 June 2024) and the resulting contigs were elongated by successive mapping of the LR, SR, and SPAdes (only for *Lgh*) and Flye assemblies. The *Lpm* mitogenome was assembled both de novo and using the *Lgh* mitogenome as a reference. For *Lgh*, as we had both LR and SR, we compared the hybrid (SPAdes), LR-only (Flye), and SR-only (Megahit) assemblies using the QUality ASsessment Tool (QUAST) [115].

CDS and tRNA annotations were performed using Geseq and BLASTx [116,117]. Rfam was used to annotate the rRNA [118]. All annotations were manually verified and biocurated. OGDRAW was used for the graphic representation of the mitogenomes [119]. For the RNA editing predictions, CDSs were extracted and Deepred-MT was used to find the C-to-U post-transcriptional modifications [120]. Five genes were only analyzed in *Lpm* (*atpF*, *ndhC*, *ndhJ*, *ndhK*, and *rps2*) as they were annotated as pseudogenes in the *Lgh* mitogenome. Mobile elements were annotated using DANTE via the RepeatExplorer Galaxy server and each region identified was checked for both gene designation and boundaries using BlastN and BlastP [121].

4.3. Sequence Analysis

REPuter on BiBiServ2 (<https://bibiserv.cebitec.uni-bielefeld.de>, accessed on 28 June 2024) identified direct and reverse repeats (maximum computed repeats = 500; minimal repeat size = 8) [122]. Simple sequence repeats (SSR) were annotated by MISA with the parameters set to 10, 5, 4, 3, 3, and 3 for the mono-, di-, tri-, tetra-, penta-, and hexanucleotides, respectively [123]. Tandem Repeats Finder was used to identify tandem repeats [124]. Chloroplast sequence insertions in the mitochondrial genome were found using ProgressiveMauve with the plastome [125]. Circular repeats maps were made using shinyCircos [126]. Each repeat longer than 1 kb was extracted, along with 1 kb regions on each side, and recombination events were constructed. Then, the LR were mapped to all of these sequences using the 'Map to Reference' function in Geneious, and sequences that overlapped the repeats and covered at least 100 bp of the two flanking 1 kb regions were tallied for all conformations and utilized to identify potential recombinations. If the repeats were in the same direction, the molecule was divided into two molecules at the repeat locations. If the repeats were in the opposite direction, an inversion occurred between the

two repeats. A general sequences comparison of the *Ludwigia* mitogenomes was conducted by using LastZ [127]. ProgressiveMauve identified LCBs (locally collinear blocks) between the *Lgh* and *Lpm* mitogenomes.

4.4. Codon Usage Bias and Ka/Ks Calculation

All CDSs were extracted from each *Ludwigia* mitogenome and concatenate. MEGA was used to count the presence of each codon and calculate their RSCU (relative synonymous codon usage) [126].

The complete mitogenomes of 13 species belonging to the Myrtales order were collected from the NCBI RefSeq database: *Oenothera biennis* (MZ934756), *Oenothera elata* (MZ934757), *Oenothera villaricae* (MZ934755) [37], *Chamaenerion angustifolium* (OX328283), *Lagerstroemia indica* (NC 035616), *Eucalyptus grandis* (NC 040010) [91], *Eucalyptus urophylla* × *grandis* (OQ947727), *Melastoma dodecandrum* (OR148386) [94], *Melastoma candidum* (NC 071383), *Melastoma sanguineum* (NC 071384), *Punica granatum* (NC 071229) [89], *Rhodomyrtus tomentosa* (NC 071968) [90], and *Syzygium samarangense* (NC 079700) [107]. In addition to these 13 mitogenomes, the mitogenome of *Geranium maderense* (NC 027000) from the Geraniales order was also collected and used as the outgroup species. *C. angustifolium* was annotated with Geseq [116]. Thirty-one genes present in at least 14 of the 16 studied mitogenomes (14 collected in database and *Lpm* and *Lgh*) were selected (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *nad9*, *sdh4*, *cob*, *cox1*, *cox2*, *cox3*, *atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *ccmB*, *ccmC*, *ccmFc*, *matR*, *mttb*, *rpl2*, *rpl5*, *rpl10*, *rps1*, *rps3*, *rps4*, and *rps14*). Protein coding genes were aligned using MAFFT, and MEGA was used to calculate Ka (non-synonymous substitution rate) and Ks (synonymous substitution rate) with the Nei–Gojobori (Jukes–Cantor) method, of these 31 genes between *Lgh* as the reference, and the other 15 species [127,128].

4.5. Molecular Mass, Isoelectric Point and Subcellular Localization Predictions

Sequences of mitochondrial protein coding genes were translated by Geneious. Protein sequences were used to calculate the molecular mass and isoelectric point with Sequence Manipulation Suite [129]. These were also used for subcellular localization prediction with Plant-mPLoc [130].

4.6. Phylogenetic and Comparative Analysis

Mitogenomes used in the Ka/Ks calculation were also used for phylogenetic analysis. Sixteen genes (*nad1*, *nad2*, *nad4*, *nad5*, *nad6*, *nad9*, *cob*, *cox1*, *cox3*, *atp4*, *atp6*, *atp8*, *ccmB*, *ccmC*, *ccmFc*, and *rpl10*), common between the 16 mitogenomes selected, were concatenated and aligned with MAFFT. MEGA was used to make a tree with the maximum likelihood method using a bootstrap of 1000 r.

5. Conclusions

To conclude, we obtained the first high-quality mitogenomes of two *Ludwigia* species, *Ludwigia peploides* subsp. *montevidensis* and *Ludwigia grandiflora* subsp. *hexapetala*. The *Lpm* and *Lgh* mitogenomes were the first of the Ludwigioideae subfamily to be sequenced, which regroups 83 species, and the fifth and sixth in the Onagraceae family, which includes approximately 660 species, and correspond to the eighteenth and nineteenth to be sequenced in the Myrtales order. The fact that there are so few mitogenomes in the Myrtales order compared to plastomes (342 in the RefSeq database in April 2024) highlights how challenging it is to assemble mitogenomes. In addition, *Lgh* and *Lpm* mitogenomes were the first to be found as two circular molecules in the Myrtales order. Our results can be used as a reference for the enrichment of mitochondrial resources in the Myrtales order, which is considered to be one of the largest in the Angiosperms with around 12,000 species [131].

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25137283/s1>.

Author Contributions: D.B. and F.B.-H. designed the project and defined the methodological approaches. G.D. carried out the data analysis and tested all of the data assembly and biocuration strategies. All authors wrote different parts of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by INRAE through the doctoral grant of Guillaume Doré.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contribution presented in the study are included in this article (see Supplementary Materials). Accession numbers of the *Lpm* and *Lgh* mitogenomes are indicated in the main manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Rai, P.K. Environmental Degradation by Invasive Alien Plants in the Anthropocene: Challenges and Prospects for Sustainable Restoration. *Anthr. Sci.* **2022**, *1*, 5–28. [CrossRef]
- Pyšek, P.; Hulme, P.E.; Simberloff, D.; Bacher, S.; Blackburn, T.M.; Carlton, J.T.; Dawson, W.; Essl, F.; Foxcroft, L.C.; Genovesi, P.; et al. Scientists' Warning on Invasive Alien Species. *Biol. Rev.* **2020**, *95*, 1511–1534. [CrossRef]
- Gioria, M.; Hulme, P.E.; Richardson, D.M.; Pyšek, P. Why Are Invasive Plants Successful? *Annu. Rev. Plant Biol.* **2023**, *74*, 635–670. [CrossRef] [PubMed]
- Mikulyuk, A. *Ludwigia grandiflora* (Water Primrose). *CABI Compend.* **2009**, 109148. [CrossRef]
- Grewell, B.; Netherland, M.; Skaer Thomason, M. Establishing Research and Management Priorities for Invasive Water Primroses (*Ludwigia* spp.). 2016. Available online: https://www.researchgate.net/publication/295856718_Establishing_Research_and_Management_Priorities_for_Invasive_Water_Primroses_Ludwigia_spp (accessed on 28 June 2024).
- Okada, M.; Grewell, B.J.; Jasieniuk, M. Clonal Spread of Invasive *Ludwigia hexapetala* and *L. Grandiflora* in Freshwater Wetlands of California. *Aquat. Bot.* **2009**, *91*, 123–129. [CrossRef]
- Thouvenot, L.; Haury, J.; Thiebaut, G. A Success Story: Water Primroses, Aquatic Plant Pests. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **2013**, *23*, 790–803. [CrossRef]
- Dandelot, S.; Matheron, R.; Le Petit, J.; Verlaque, R.; Cazaubon, A. Variations Temporelles des Paramètres Physicochimiques et Microbiologiques de Trois Écosystèmes Aquatiques (Sud-est de la France) Envahis par des *Ludwigia*. *Comptes Rendus Biol.* **2005**, *328*, 991–999. [CrossRef]
- Portillo Lemus, L.O.; Bozec, M.; Harang, M.; Coudreuse, J.; Haury, J.; Stoeckel, S.; Barloy, D. Self-Incompatibility Limits Sexual Reproduction Rather than Environmental Conditions in an Invasive Water Primrose. *Plant-Environ. Interact.* **2021**, *2*, 74–86. [CrossRef]
- Thiebaut, G.; Dutartre, A. Management of invasive aquatic plants in France. In *Aquatic Ecosystem Research Trends*; Nova Science Pub Inc.: New York, NY, USA, 2009.
- Nehring, S.; Kolthoff, D. The Invasive Water Primrose *Ludwigia grandiflora* (Michaux) Greuter & Burdet (Spermatophyta: Onagraceae) in Germany: First Record and Ecological Risk Assessment. *Aquat. Invasions* **2011**, *6*, 83–89. [CrossRef]
- Celesti-Grapow, L.; Alessandrini, A.; Arrigoni, P.V.; Banfi, E.; Bernardo, L.; Bovio, M.; Brundu, G.; Cagiotti, M.R.; Camarda, I.; Carli, E.; et al. Inventory of the Non-native Flora of Italy. *Plant Biosyst.-Int. J. Deal. All Asp. Plant Biol.* **2009**, *143*, 386–430. [CrossRef]
- Bou Manobens, J.; Portillo, L.; Curcó i Massip, A. New Contributions to Allochthonous *Ludwigia* Species (Onagraceae) on Catalonia. *Butlletí Inst. Catalana D'història Nat.* **2019**, 45–48. [CrossRef]
- Lindemann-Matthies, P. Beasts or Beauties? Laypersons' Perception of Invasive Alien Plant Species in Switzerland and Attitudes towards Their Management. *NeoBiota* **2016**, *29*, 15–33. [CrossRef]
- Gallardo, B.; Aldridge, D.C. The 'Dirty Dozen': Socio-Economic Factors Amplify the Invasion Potential of 12 High-Risk Aquatic Invasive Species in Great Britain and Ireland. *J. Appl. Ecol.* **2013**, *50*, 757–766. [CrossRef]
- Arslan, Z.F.; Uludag, A.; Uremis, I. Status of Invasive Alien Plants Included in EPPO Lists in Turkey. *EPPO Bull.* **2015**, *45*, 66–72. [CrossRef]
- Hieda, S.; Kaneko, Y.; Nakagawa, M.; Noma, N. *Ludwigia grandiflora* (Michx.) Greuter & Burdet subsp. *Hexapetala* (Hook. & Arn.) GL Nesom & Kartesz, an Invasive Aquatic Plant in Lake Biwa, the Largest Lake in Japan 2020. *Acta Phytotaxon. Geobot.* **2020**, *71*, 65–71.
- EPPO. *Ludwigia grandiflora* and *L. Peploides* onagraceae—Water Primroses. *EPPO Bull.* **2011**, *41*, 414–418. [CrossRef]
- Sarika, M. *Ludwigia peploides* subsp. *Montevidensis*, a New Alien Taxon for the Flora of Greece and the Balkans. *J. Biol. Res.* **2006**, *5*, 71–78.
- Dandelot, S. Les *Ludwigia* spp. invasives du Sud de la France: Historique, Biosystématique, Biologie et Ecologie. Ph.D. Thesis, Université Paul Cézanne: Aix-Marseille III, Marseille, France, 2004.

21. Muller, S. *Plantes invasives en France*; Muséum National d'Histoire Naturelle: Paris, France, 2004.
22. Rolon, A.S.; Lacerda, T.; Maltchik, L.; Guadagnin, D.L. Influence of Area, Habitat and Water Chemistry on Richness and Composition of Macrophyte Assemblages in Southern Brazilian Wetlands. *J. Veg. Sci.* **2008**, *19*, 221–228. [[CrossRef](#)]
23. Hernández-R, J.; Rangel-Ch, J.O. Vegetation of the Wetland Jaboque (Bogotá, DC). *Caldasia* **2009**, *31*, 355–379.
24. Haury, J.; Druel, A.; Cabral, T.; Paulet, Y.; Bozec, M.; Coudreuse, J. Which Adaptations of Some Invasive *Ludwigia* Spp. (Rosidae, Onagraceae) Populations Occur in Contrasting Hydrological Conditions in Western France? *Hydrobiologia* **2014**, *737*, 45–56. [[CrossRef](#)]
25. Billet, K.; Genitoni, J.; Bozec, M.; Renault, D.; Barloy, D. Aquatic and Terrestrial Morphotypes of the Aquatic Invasive Plant, *Ludwigia grandiflora*, Show Distinct Morphological and Metabolomic Responses. *Ecol. Evol.* **2018**, *8*, 2568–2579. [[CrossRef](#)] [[PubMed](#)]
26. Moravcová, L.; Pyšek, P.; Jarošík, V.; Pergl, J. Getting the Right Traits: Reproductive and Dispersal Characteristics Predict the Invasiveness of Herbaceous Plant Species. *PLoS ONE* **2015**, *10*, e0123634. [[CrossRef](#)] [[PubMed](#)]
27. Mounger, J.; Ainouche, M.L.; Bossdorf, O.; Cavé-Radet, A.; Li, B.; Parepa, M.; Salmon, A.; Yang, J.; Richards, C.L. Epigenetics and the Success of Invasive Plants. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2021**, *376*, 20200117. [[CrossRef](#)] [[PubMed](#)]
28. Hodgkinson, T.R.; Perdereau, A.; Klaas, M.; Cormican, P.; Barth, S. Genotyping by Sequencing and Plastome Analysis Finds High Genetic Variability and Geographical Structure in *Dactylis glomerata* L. in Northwest Europe Despite Lack of Ploidy Variation. *Agronomy* **2019**, *9*, 342. [[CrossRef](#)]
29. Chown, S.L.; Hodgins, K.A.; Griffin, P.C.; Oakeshott, J.G.; Byrne, M.; Hoffmann, A.A. Biological Invasions, Climate Change and Genomics. *Evol. Appl.* **2015**, *8*, 23–46. [[CrossRef](#)]
30. Hamelin, R.; Roe, A. Genomic Biosurveillance of Forest Invasive Alien Enemies: A Story Written in Code. *Evol. Appl.* **2019**, *13*, 12853. [[CrossRef](#)] [[PubMed](#)]
31. Liu, S.-H.; Edwards, C.; Hoch, P.C.; Raven, P.H.; Barber, J.C. Complete Plastome Sequence of *Ludwigia octovalvis* (Onagraceae), a Globally Distributed Wetland Plant. *Genome Announc.* **2016**, *4*, e01274-16. [[CrossRef](#)]
32. Barloy-Hubler, F.; Gac, A.-L.L.; Boury, C.; Guichoux, E.; Barloy, D. Sequencing, de Novo Assembly of *Ludwigia* Plastomes, and Comparative Analysis within the Onagraceae Family. *bioRxiv* **2023**. [[CrossRef](#)]
33. Barloy, D.; Portillo-Lemus, L.; Krueger-Hadfield, S.; Huteau, V.; Coriton, O. Genomic relationships among diploid and polyploid species of the genus *Ludwigia* L. section *Jussiaea* using a combination of molecular cytogenetic, morphological, and crossing investigations. *Peer Community J.* **2024**, *4*, e8. [[CrossRef](#)]
34. Wang, Y.; Yu, J.; Jiang, M.; Lei, W.; Zhang, X.; Tang, H. Sequencing and Assembly of Polyploid Genomes. *Methods Mol. Biol.* **2023**, *2545*, 429–458. [[CrossRef](#)]
35. Jacoby, R.P.; Li, L.; Huang, S.; Pong Lee, C.; Millar, A.H.; Taylor, N.L. Mitochondrial Composition, Function and Stress Response in Plants. *J. Integr. Plant Biol.* **2012**, *54*, 887–906. [[CrossRef](#)]
36. Mackenzie, S.; McIntosh, L. Higher Plant Mitochondria. *Plant Cell* **1999**, *11*, 571–585. [[CrossRef](#)] [[PubMed](#)]
37. Fischer, A.; Dotzek, J.; Walther, D.; Greiner, S. Graph-Based Models of the Oenothera Mitochondrial Genome Capture the Enormous Complexity of Higher Plant Mitochondrial DNA Organization. *NAR Genom. Bioinform.* **2022**, *4*, lqac027. [[CrossRef](#)]
38. Kozik, A.; Rowan, B.A.; Lavelle, D.; Berke, L.; Schranz, M.E.; Michelmore, R.W.; Christensen, A.C. The Alternative Reality of Plant Mitochondrial DNA: One Ring Does Not Rule Them All. *PLoS Genet.* **2019**, *15*, e1008373. [[CrossRef](#)]
39. Liu, D.; Zeng, M.; Wu, Y.; Du, Y.; Liu, J.; Luo, S.; Zeng, Y. Comparative Transcriptomic Analysis Provides Insights into the Molecular Basis Underlying Pre-Harvest Sprouting in Rice. *BMC Genom.* **2022**, *23*, 771. [[CrossRef](#)] [[PubMed](#)]
40. Varré, J.-S.; D'Agostino, N.; Touzet, P.; Gallina, S.; Tamburino, R.; Cantarella, C.; Ubrig, E.; Cardi, T.; Drouard, L.; Gualberto, J.M.; et al. Complete Sequence, Multichromosomal Architecture and Transcriptome Analysis of the *Solanum tuberosum* Mitochondrial Genome. *Int. J. Mol. Sci.* **2019**, *20*, 4788. [[CrossRef](#)]
41. Li, J.; Xu, Y.; Shan, Y.; Pei, X.; Yong, S.; Liu, C.; Yu, J. Assembly of the Complete Mitochondrial Genome of an Endemic Plant, *Scutellaria tsinyunensis*, Revealed the Existence of Two Conformations Generated by a Repeat-Mediated Recombination. *Planta* **2021**, *254*, 36. [[CrossRef](#)] [[PubMed](#)]
42. Tsujimura, M.; Kaneko, T.; Sakamoto, T.; Kimura, S.; Shigyo, M.; Yamagishi, H.; Terachi, T. Multichromosomal Structure of the Onion Mitochondrial Genome and a Transcript Analysis. *Mitochondrion* **2019**, *46*, 179–186. [[CrossRef](#)] [[PubMed](#)]
43. Lloyd Evans, D.; Hlongwane, T.T.; Joshi, S.V.; Riaño Pachón, D.M. The Sugarcane Mitochondrial Genome: Assembly, Phylogenetics and Transcriptomics. *PeerJ* **2019**, *7*, e7558. [[CrossRef](#)]
44. Kim, C.-K.; Kim, Y.-K. The Multipartite Mitochondrial Genome of *Fallopia multiflora* (Caryophyllales: Polygonaceae). *Mitochondrial DNA B Resour.* **2018**, *3*, 155–156. [[CrossRef](#)]
45. Yang, H.; Chen, H.; Ni, Y.; Li, J.; Cai, Y.; Wang, J.; Liu, C. Mitochondrial Genome Sequence of *Salvia officinalis* (Lamiales: Lamiaceae) Suggests Diverse Genome Structures in Cogeneric Species and Finds the Stop Gain of Genes through RNA Editing Events. *Int. J. Mol. Sci.* **2023**, *24*, 5372. [[CrossRef](#)]
46. Guo, Y.; Li, Z.; Jin, S.; Chen, S.; Li, F.; Wu, H. Assembly and Comparative Analysis of the Complete Mitochondrial Genome of Two Species of Calla Lilies (*Zantedeschia*, Araceae). *Int. J. Mol. Sci.* **2023**, *24*, 9566. [[CrossRef](#)]
47. Kang, S.-H.; Oh, J.-H.; Kim, H.-J.; Kim, C.-K. The Multipartite Mitochondrial Genome of *Cynanchum wilfordii* (Gentianales: Apocynaceae). *Mitochondrial DNA B Resour.* **2017**, *2*, 720–721. [[CrossRef](#)]

48. Chang, S.; Yang, T.; Du, T.; Huang, Y.; Chen, J.; Yan, J.; He, J.; Guan, R. Mitochondrial Genome Sequencing Helps Show the Evolutionary Mechanism of Mitochondrial Genome Formation in Brassica. *BMC Genom.* **2011**, *12*, 497. [[CrossRef](#)]
49. Alverson, A.J.; Rice, D.W.; Dickinson, S.; Barry, K.; Palmer, J.D. Origins and Recombination of the Bacterial-Sized Multichromosomal Mitochondrial Genome of Cucumber. *Plant Cell* **2011**, *23*, 2499–2513. [[CrossRef](#)]
50. Rice, D.W.; Alverson, A.J.; Richardson, A.O.; Young, G.J.; Sanchez-Puerta, M.V.; Munzinger, J.; Barry, K.; Boore, J.L.; Zhang, Y.; dePamphilis, C.W.; et al. Horizontal Transfer of Entire Genomes via Mitochondrial Fusion in the Angiosperm *Amborella*. *Science* **2013**, *342*, 1468–1473. [[CrossRef](#)]
51. Yang, J.-X.; Dierckxsens, N.; Bai, M.-Z.; Guo, Y.-Y. Multichromosomal Mitochondrial Genome of *Paphiopedilum Micranthum*: Compact and Fragmented Genome, and Rampant Intracellular Gene Transfer. *Int. J. Mol. Sci.* **2023**, *24*, 3976. [[CrossRef](#)]
52. Sanchez-Puerta, M.V.; García, L.E.; Wohlfeiler, J.; Ceriotti, L.F. Unparalleled Replacement of Native Mitochondrial Genes by Foreign Homologs in a Holoparasitic Plant. *New Phytol.* **2017**, *214*, 376–387. [[CrossRef](#)]
53. Wu, Z.; Cuthbert, J.M.; Taylor, D.R.; Sloan, D.B. The Massive Mitochondrial Genome of the Angiosperm *Silene Noctiflora* Is Evolving by Gain or Loss of Entire Chromosomes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10185–10191. [[CrossRef](#)]
54. Sloan, D.B.; Alverson, A.J.; Wu, M.; Palmer, J.D.; Taylor, D.R. Recent Acceleration of Plastid Sequence and Structural Evolution Coincides with Extreme Mitochondrial Divergence in the Angiosperm Genus *Silene*. *Genome Biol. Evol.* **2012**, *4*, 294–306. [[CrossRef](#)]
55. Skippington, E.; Barkman, T.J.; Rice, D.W.; Palmer, J.D. Miniaturized Mitogenome of the Parasitic Plant *Viscum Scurreloideum* Is Extremely Divergent and Dynamic and Has Lost All Nad Genes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E3515–E3524. [[CrossRef](#)] [[PubMed](#)]
56. Putintseva, Y.A.; Bondar, E.I.; Simonov, E.P.; Sharov, V.V.; Oreshkova, N.V.; Kuzmin, D.A.; Konstantinov, Y.M.; Shmakov, V.N.; Belkov, V.I.; Sadovsky, M.G.; et al. *Siberian larch* (*Larix sibirica* Ledeb.) mitochondrial genome assembled using both short and long nucleotide sequence reads is currently the largest known mitogenome. *BMC Genom.* **2020**, *21*, 654. [[CrossRef](#)] [[PubMed](#)]
57. Mower, J.P.; Sloan, D.B.; Alverson, A.J. Plant Mitochondrial Genome Diversity: The Genomics Revolution. In *Plant Genome; Wendel, J.F., Greilhuber, J., Dolezel, J., Leitch, I.J., Eds.; Springer: Vienna, Austria, 2012; Volume 1, pp. 123–144, ISBN 978-3-7091-1129-1.*
58. Wynn, E.L.; Christensen, A.C. Repeats of Unusual Size in Plant Mitochondrial Genomes: Identification, Incidence and Evolution. *G3 Genes Genomes Genet.* **2019**, *9*, 549–559. [[CrossRef](#)] [[PubMed](#)]
59. Bruenn, J.A.; Warner, B.E.; Yerramsetty, P. Widespread Mitovirus Sequences in Plant Genomes. *PeerJ* **2015**, *3*, e876. [[CrossRef](#)] [[PubMed](#)]
60. Choi, I.-S.; Wojciechowski, M.F.; Ruhlman, T.A.; Jansen, R.K. In and out: Evolution of Viral Sequences in the Mitochondrial Genomes of Legumes (Fabaceae). *Mol. Phylogenet. Evol.* **2021**, *163*, 107236. [[CrossRef](#)]
61. Knoop, V.; Unseld, M.; Marienfeld, J.; Brandt, P.; Sünkel, S.; Ullrich, H.; Brennicke, A. Copia-, Gypsy- and LINE-like Retrotransposon Fragments in the Mitochondrial Genome of *Arabidopsis thaliana*. *Genetics* **1996**, *142*, 579–585. [[CrossRef](#)] [[PubMed](#)]
62. Yu, X.; Jiang, W.; Tan, W.; Zhang, X.; Tian, X. Deciphering the Organelle Genomes and Transcriptomes of a Common Ornamental Plant *Ligustrum Quihoui* Reveals Multiple Fragments of Transposable Elements in the Mitogenome. *Int. J. Biol. Macromol.* **2020**, *165*, 1988–1999. [[CrossRef](#)]
63. de Abreu, V.A.C.; Moysés Alves, R.; Silva, S.R.; Ferro, J.A.; Domingues, D.S.; Miranda, V.F.O.; Varani, A.M. Comparative Analyses of *Theobroma Cacao* and *T. Grandiflorum* Mitogenomes Reveal Conserved Gene Content Embedded within Complex and Plastic Structures. *Gene* **2023**, *849*, 146904. [[CrossRef](#)]
64. Sloan, D.B. One Ring to Rule Them All? Genome Sequencing Provides New Insights into the ‘Master Circle’ Model of Plant Mitochondrial DNA Structure. *New Phytol.* **2013**, *200*, 978–985. [[CrossRef](#)]
65. Arrieta-Montiel, M.P.; Mackenzie, S.A. Plant Mitochondrial Genomes and Recombination. In *Plant Mitochondria; Kempken, F., Ed.; Springer: New York, NY, USA, 2011; pp. 65–82. ISBN 978-0-387-89781-3.*
66. Zhang, X.; Shan, Y.; Li, J.; Qin, Q.; Yu, J.; Deng, H. Assembly of the Complete Mitochondrial Genome of *Pereskia Aculeata* Revealed That Two Pairs of Repetitive Elements Mediated the Recombination of the Genome. *Int. J. Mol. Sci.* **2023**, *24*, 8366. [[CrossRef](#)]
67. Yang, Z.; Ni, Y.; Lin, Z.; Yang, L.; Chen, G.; Nijjati, N.; Hu, Y.; Chen, X. De Novo Assembly of the Complete Mitochondrial Genome of Sweet Potato (*Ipomoea batatas* [L.] Lam) Revealed the Existence of Homologous Conformations Generated by the Repeat-Mediated Recombination. *BMC Plant Biol.* **2022**, *22*, 285. [[CrossRef](#)]
68. Sullivan, A.R.; Eldfjell, Y.; Schiffthaler, B.; Delhomme, N.; Asp, T.; Hebelstrup, K.H.; Keech, O.; Öberg, L.; Møller, I.M.; Arvestad, L.; et al. The Mitogenome of Norway Spruce and a Reappraisal of Mitochondrial Recombination in Plants. *Genome Biol. Evol.* **2020**, *12*, 3586–3598. [[CrossRef](#)]
69. Lonsdale, D.M.; Brears, T.; Hodge, T.P.; Melville, S.E.; Rottmann, W.H.; Leaver, C.J.; Lonsdale, D.M. The Plant Mitochondrial Genome: Homologous Recombination as a Mechanism for Generating Heterogeneity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1997**, *319*, 149–163. [[CrossRef](#)]
70. Gualberto, J.M.; Newton, K.J. Plant Mitochondrial Genomes: Dynamics and Mechanisms of Mutation. *Annu. Rev. Plant Biol.* **2017**, *68*, 225–252. [[CrossRef](#)]
71. Kubo, T.; Newton, K.J. Angiosperm Mitochondrial Genomes and Mutations. *Mitochondrion* **2008**, *8*, 5–14. [[CrossRef](#)]

72. Marienfeld, J.; Unseld, M.; Brennicke, A. The Mitochondrial Genome of Arabidopsis Is Composed of Both Native and Immigrant Information. *Trends Plant Sci.* **1999**, *4*, 495–502. [[CrossRef](#)]
73. Hao, W.; Palmer, J.D. Fine-Scale Mergers of Chloroplast and Mitochondrial Genes Create Functional, Transcompartmentally Chimeric Mitochondrial Genes. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16728–16733. [[CrossRef](#)]
74. Marechal-Drouard, L.; Weil, J.H.; Dietrich, A. Transfer RNAs and Transfer RNA Genes in Plants. *Annu. Rev. Plant Biol.* **1993**, *44*, 13–32. [[CrossRef](#)]
75. Nakazono, M.; Nishiwaki, S.; Tsutsumi, N.; Hirai, A. A Chloroplast-Derived Sequence Is Utilized as a Source of Promoter Sequences for the Gene for Subunit 9 of NADH Dehydrogenase (Nad9) in Rice Mitochondria. *Mol. Gen. Genet.* **1996**, *252*, 371–378. [[CrossRef](#)]
76. Wang, D.; Rousseau-Gueutin, M.; Timmis, J.N. Plastid Sequences Contribute to Some Plant Mitochondrial Genes. *Mol. Biol. Evol.* **2012**, *29*, 1707–1711. [[CrossRef](#)] [[PubMed](#)]
77. Fox, T.D.; Leaver, C.J. The Zea Mays Mitochondrial Gene Coding Cytochrome Oxidase Subunit II Has an Intervening Sequence and Does Not Contain TGA Codons. *Cell* **1981**, *26*, 315–323. [[CrossRef](#)] [[PubMed](#)]
78. Chao, S.; Sederoff, R.R.; Levings, C.S. Partial Sequence Analysis of the 5S to 18S rRNA Gene Region of the Maize Mitochondrial Genome. *Plant Physiol.* **1983**, *71*, 190–193. [[CrossRef](#)]
79. Sloan, D.B.; Alverson, A.J.; Chuckalovcak, J.P.; Wu, M.; McCauley, D.E.; Palmer, J.D.; Taylor, D.R. Rapid Evolution of Enormous, Multichromosomal Genomes in Flowering Plant Mitochondria with Exceptionally High Mutation Rates. *PLoS Biol.* **2012**, *10*, e1001241. [[CrossRef](#)]
80. Knoop, V. The Mitochondrial DNA of Land Plants: Peculiarities in Phylogenetic Perspective. *Curr. Genet.* **2004**, *46*, 123–139. [[CrossRef](#)] [[PubMed](#)]
81. Unseld, M.; Marienfeld, J.R.; Brandt, P.; Brennicke, A. The Mitochondrial Genome of Arabidopsis Thaliana Contains 57 Genes in 366,924 Nucleotides. *Nat. Genet.* **1997**, *15*, 57–61. [[CrossRef](#)]
82. Adams, K.L.; Palmer, J.D. Evolution of Mitochondrial Gene Content: Gene Loss and Transfer to the Nucleus. *Mol. Phylogenet. Evol.* **2003**, *29*, 380–395. [[CrossRef](#)] [[PubMed](#)]
83. Ichinose, M.; Sugita, M. RNA Editing and Its Molecular Mechanism in Plant Organelles. *Genes* **2016**, *8*, 5. [[CrossRef](#)] [[PubMed](#)]
84. Small, I.D.; Schallenberg-Rüdinger, M.; Takenaka, M.; Mireau, H.; Ostersetzer-Biran, O. Plant Organellar RNA Editing: What 30 Years of Research Has Revealed. *Plant J.* **2020**, *101*, 1040–1056. [[CrossRef](#)]
85. Gualberto, J.M.; Mileshina, D.; Wallet, C.; Niazi, A.K.; Weber-Lotfi, F.; Dietrich, A. The Plant Mitochondrial Genome: Dynamics and Maintenance. *Biochimie* **2014**, *100*, 107–120. [[CrossRef](#)]
86. Wang, S.; Song, Q.; Li, S.; Hu, Z.; Dong, G.; Song, C.; Huang, H.; Liu, Y. Assembly of a Complete Mitogenome of Chrysanthemum Nankingense Using Oxford Nanopore Long Reads and the Diversity and Evolution of Asteraceae mitogenomes. *Genes* **2018**, *9*, 547. [[CrossRef](#)]
87. Kovar, L.; Nageswara-Rao, M.; Ortega-Rodriguez, S.; Dugas, D.V.; Straub, S.; Cronn, R.; Strickler, S.R.; Hughes, C.E.; Hanley, K.A.; Rodriguez, D.N.; et al. PacBio-Based Mitochondrial Genome Assembly of *Leucaena trichandra* (Leguminosae) and an Intrageneric Assessment of Mitochondrial RNA Editing. *Genome Biol. Evol.* **2018**, *10*, 2501–2517. [[CrossRef](#)] [[PubMed](#)]
88. The Angiosperm Phylogeny Group; Chase, M.W.; Christenhusz, M.J.M.; Fay, M.F.; Byng, J.W.; Judd, W.S.; Soltis, D.E.; Mabberley, D.J.; Sennikov, A.N.; Soltis, P.S.; et al. An Update of the Angiosperm Phylogeny Group Classification for the Orders and Families of Flowering Plants: APG IV. *Bot. J. Linn. Soc.* **2016**, *181*, 12385. [[CrossRef](#)]
89. Lu, G.; Zhang, K.; Que, Y.; Li, Y. Assembly and Analysis of the First Complete Mitochondrial Genome of Punica Granatum and the Gene Transfer from Chloroplast Genome. *Front. Plant Sci.* **2023**, *14*, 1132551. [[CrossRef](#)] [[PubMed](#)]
90. Yu, X.; Ma, Z.; Liu, S.; Duan, Z. Analysis of the *Rhodomyrtus tomentosa* Mitochondrial Genome: Insights into Repeat-Mediated Recombination and Intra-Cellular DNA Transfer. *Gene* **2024**, *909*, 148288. [[CrossRef](#)] [[PubMed](#)]
91. Pinard, D.; Myburg, A.A.; Mizrachi, E. The Plastid and Mitochondrial Genomes of Eucalyptus Grandis. *BMC Genom.* **2019**, *20*, 132. [[CrossRef](#)] [[PubMed](#)]
92. Fukasawa, Y.; Driguez, P.; Bougouffa, S.; Carty, K.; Putra, A.; Cheung, M.-S.; Ermini, L. Plasticity of Repetitive Sequences Demonstrated by the Complete Mitochondrial Genome of Eucalyptus Camaldulensis. *Front. Plant Sci.* **2024**, *15*, 1339594. [[CrossRef](#)]
93. Wu, H.-H.; Zhao, X.-H.; Zong, X.-Y.; Ding, R.; Chen, X.-H. Complete Mitochondrial Genome of *Medinilla magnifica* (Myrtales, Melastomataceae). *Mitochondrial DNA Part B* **2020**, *5*, 1716–1717. [[CrossRef](#)]
94. Zhou, Y.; Zheng, R.; Peng, Y.; Chen, J.; Zhu, X.; Xie, K.; Ahmad, S.; Chen, J.; Wang, F.; Shen, M.; et al. The First Mitochondrial Genome of *Melastoma dodecandrum* Resolved Structure Evolution in Melastomataceae and Micro Inversions from Inner Horizontal Gene Transfer. *Ind. Crops Prod.* **2023**, *205*, 117390. [[CrossRef](#)]
95. Wu, Z.-Q.; Liao, X.-Z.; Zhang, X.-N.; Tembrock, L.R.; Broz, A. Genomic Architectural Variation of Plant Mitochondria—A Review of Multichromosomal Structuring. *J. Syst. Evol.* **2020**, *60*, 160–168. [[CrossRef](#)]
96. Cole, L.W.; Guo, W.; Mower, J.P.; Palmer, J.D. High and Variable Rates of Repeat-Mediated Mitochondrial Genome Rearrangement in a Genus of Plants. *Mol. Biol. Evol.* **2018**, *35*, 2773–2785. [[CrossRef](#)]
97. Bi, C.; Qu, Y.; Hou, J.; Wu, K.; Ye, N.; Yin, T. Deciphering the Multi-Chromosomal Mitochondrial Genome of *Populus simonii*. *Front. Plant Sci.* **2022**, *13*, 914635. [[CrossRef](#)] [[PubMed](#)]

98. Yang, H.; Chen, H.; Ni, Y.; Li, J.; Cai, Y.; Ma, B.; Yu, J.; Wang, J.; Liu, C. De Novo Hybrid Assembly of the *Salvia miltiorrhiza* Mitochondrial Genome Provides the First Evidence of the Multi-Chromosomal Mitochondrial DNA Structure of Salvia Species. *Int. J. Mol. Sci.* **2022**, *23*, 14267. [[CrossRef](#)] [[PubMed](#)]
99. Sloan, D.B.; Alverson, A.J.; Štorchová, H.; Palmer, J.D.; Taylor, D.R. Extensive Loss of Translational Genes in the Structurally Dynamic Mitochondrial Genome of the Angiosperm *Silene latifolia*. *BMC Evol. Biol.* **2010**, *10*, 274. [[CrossRef](#)] [[PubMed](#)]
100. You, C.; Cui, T.; Zhang, C.; Zang, S.; Su, Y.; Que, Y. Assembly of the Complete Mitochondrial Genome of *Gelsemium elegans* Revealed the Existence of Homologous Conformations Generated by a Repeat Mediated Recombination. *Int. J. Mol. Sci.* **2023**, *24*, 527. [[CrossRef](#)] [[PubMed](#)]
101. Wang, X.-C.; Chen, H.; Yang, D.; Liu, C. Diversity of Mitochondrial Plastid DNAs (MTPTs) in Seed Plants. *Mitochondrial DNA Part A* **2018**, *29*, 635–642. [[CrossRef](#)] [[PubMed](#)]
102. Tang, J.; Luo, Z.; Zhang, J.; Chen, L.; Li, L. Multi-Chromosomal Mitochondrial Genome of Medicinal Plant *Acorus tatarinowii* (Acoraceae): Firstly Reported from Acorales Order. *Gene* **2024**, *892*, 147847. [[CrossRef](#)]
103. Yang, L.; Liu, J.; Guo, W.; Zheng, Z.; Xu, Y.; Xia, H.; Xiao, T. Insights into the Multi-Chromosomal Mitochondrial Genome Structure of the Xero-Halophytic Plant *Haloxylon ammodendron* (C.A.Mey.) Bunge Ex Fenzl. *BMC Genom.* **2024**, *25*, 123. [[CrossRef](#)]
104. Hao, W.; Liu, G.; Wang, W.; Shen, W.; Zhao, Y.; Sun, J.; Yang, Q.; Zhang, Y.; Fan, W.; Pei, S.; et al. RNA Editing and Its Roles in Plant Organelles. *Front. Genet.* **2021**, *12*, 757109. [[CrossRef](#)]
105. Yan, J.; Zhang, Q.; Yin, P. RNA Editing Machinery in Plant Organelles. *Sci. China Life Sci.* **2018**, *61*, 162–169. [[CrossRef](#)]
106. Gommans, W.M.; Mullen, S.P.; Maas, S. RNA Editing: A Driving Force for Adaptive Evolution? *BioEssays* **2009**, *31*, 1137–1145. [[CrossRef](#)]
107. Lu, G.; Li, Q. Complete Mitochondrial Genome of *Syzygium samarangense* Reveals Genomic Recombination, Gene Transfer, and RNA Editing Events. *Front. Plant Sci.* **2024**, *14*, 1301164. [[CrossRef](#)]
108. Xiong, J.; Tao, T.; Luo, Z.; Yan, S.; Liu, Y.; Yu, X.; Liu, G.; Xia, H.; Luo, L. RNA Editing Responses to Oxidative Stress between a Wild Abortive Type Male-Sterile Line and Its Maintainer Line. *Front. Plant Sci.* **2017**, *8*, 2023. [[CrossRef](#)]
109. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [[CrossRef](#)]
110. Bushnell, B.; Rood, J.; Singer, E. BBMerge—Accurate Paired Shotgun Read Merging via Overlap. *PLoS ONE* **2017**, *12*, e0185056. [[CrossRef](#)]
111. Pribelski, A.; Antipov, D.; Meleshko, D.; Lapidus, A.; Korobeynikov, A. Using SPAdes De Novo Assembler. *Curr. Protoc. Bioinform.* **2020**, *70*, e102. [[CrossRef](#)]
112. Holley, G.; Beyter, D.; Ingimundardottir, H.; Møller, P.L.; Kristmundsdottir, S.; Eggertsson, H.P.; Halldorsson, B.V. Ratatosk: Hybrid Error Correction of Long Reads Enables Accurate Variant Calling and Assembly. *Genome Biol.* **2021**, *22*, 28. [[CrossRef](#)]
113. Kolmogorov, M.; Yuan, J.; Lin, Y.; Pevzner, P.A. Assembly of Long, Error-Prone Reads Using Repeat Graphs. *Nat. Biotechnol.* **2019**, *37*, 540–546. [[CrossRef](#)]
114. Li, D.; Liu, C.-M.; Luo, R.; Sadakane, K.; Lam, T.-W. MEGAHIT: An Ultra-Fast Single-Node Solution for Large and Complex Metagenomics Assembly via Succinct de Bruijn Graph. *Bioinformatics* **2015**, *31*, 1674–1676. [[CrossRef](#)]
115. Gurevich, A.; Saveliev, V.; Vyahhi, N.; Tesler, G. QUAST: Quality Assessment Tool for Genome Assemblies. *Bioinformatics* **2013**, *29*, 1072–1075. [[CrossRef](#)]
116. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq—Versatile and Accurate Annotation of Organelle Genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [[CrossRef](#)]
117. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)] [[PubMed](#)]
118. 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*, W59–W64. [[CrossRef](#)] [[PubMed](#)]
119. Edera, A.A.; Small, I.; Milone, D.H.; Sanchez-Puerta, M.V. Deepred-Mt: Deep Representation Learning for Predicting C-to-U RNA Editing in Plant Mitochondria. *Comput. Biol. Med.* **2021**, *136*, 104682. [[CrossRef](#)] [[PubMed](#)]
120. Novák, P.; Neumann, P.; Pech, J.; Steinhaisl, J.; Macas, J. RepeatExplorer: A Galaxy-Based Web Server for Genome-Wide Characterization of Eukaryotic Repetitive Elements from next-Generation Sequence Reads. *Bioinformatics* **2013**, *29*, 792–793. [[CrossRef](#)]
121. Kurtz, S.; Schleiermacher, C. REPuter: Fast Computation of Maximal Repeats in Complete Genomes. *Bioinformatics* **1999**, *15*, 426–427. [[CrossRef](#)] [[PubMed](#)]
122. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-Web: A Web Server for Microsatellite Prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [[CrossRef](#)]
123. Benson, G. Tandem Repeats Finder: A Program to Analyze DNA Sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [[CrossRef](#)]
124. Darling, A.E.; Mau, B.; Perna, N.T. progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *PLoS ONE* **2010**, *5*, e11147. [[CrossRef](#)]
125. Yu, Y.; Ouyang, Y.; Yao, W. shinyCircos: An R/Shiny Application for Interactive Creation of Circos Plot. *Bioinformatics* **2018**, *34*, 1229–1231. [[CrossRef](#)]
126. Harris, R.S. *Improved Pairwise Alignment of Genomic DNA*; The Pennsylvania State University: University Park, PA, USA, 2007.

127. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)]
128. Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)]
129. Stothard, P. The Sequence Manipulation Suite: JavaScript Programs for Analyzing and Formatting Protein and DNA Sequences. *BioTechniques* **2000**, *28*, 1102–1104. [[CrossRef](#)] [[PubMed](#)]
130. Chou, K.-C.; Shen, H.-B. Plant-mPLoc: A Top-Down Strategy to Augment the Power for Predicting Plant Protein Subcellular Localization. *PLoS ONE* **2010**, *5*, e11335. [[CrossRef](#)]
131. Kriebel, R.; Khabbazian, M.; Sytsma, K.J. A Continuous Morphological Approach to Study the Evolution of Pollen in a Phylogenetic Context: An Example with the Order Myrtales. *PLoS ONE* **2017**, *12*, e0187228. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.