Mitochondrial Genome Analysis of Wild Rice (<i>Oryza minuta</i>) and Its Comparison with Other Related Species

Sajjad Asaf¹, Abdul Latif Khan², Abdur Rahim Khan³, Muhammad Waqas¹,³, Sang-Mo Kang¹, Muhammad Aaqil Khan¹, Raheem Shazad¹, Chang-Woo Seo¹, Jae-Ho Shin¹, In-Jung Lee¹*  

¹ School of Applied Biosciences, Kyungpook National University, Daegu, 41566, Republic of Korea, ² Chair of Oman’s Medicinal Plants & Marine Natural Products, University of Nizwa, Nizwa, 616, Oman, ³ Department of Agriculture, Abdul Wali Khan University Mardan, Mardan, Pakistan  

* ijlee@knu.ac.kr

Abstract

<i>Oryza minuta</i> (Poaceae family) is a tetraploid wild relative of cultivated rice with a BBCC genome. <i>O. minuta</i> has the potential to resist against various pathogenic diseases such as bacterial blight (BB), white backed planthopper (WBPH) and brown plant hopper (BPH). Here, we sequenced and annotated the complete mitochondrial genome of <i>O. minuta</i>. The mtDNA genome is 515,022 bp, containing 60 protein coding genes, 31 tRNA genes and two rRNA genes. The mitochondrial genome organization and the gene content at the nucleotide level are highly similar (89%) to that of <i>O. rufipogon</i>. Comparison with other related species revealed that most of the genes with known function are conserved among the Poaceae members. Similarly, <i>O. minuta</i> mt genome shared 24 protein-coding genes, 15 tRNA genes and 1 ribosomal RNA gene with other rice species (<i>indica</i> and <i>japonica</i>). The evolutionary relationship and phylogenetic analysis revealed that <i>O. minuta</i> is more closely related to <i>O. rufipogon</i> than to any other related species. Such studies are essential to understand the evolutionary divergence among species and analyze common gene pools to combat risks in the current scenario of a changing environment.

Introduction

In recent years, we have noticed a significant increase in the sequencing of organelle genomes, particularly those of economically important crop plants. To date, 300 mitochondrial (mt) and 342 complete chloroplast (cp) genomes have been submitted to GenBank Organelle Genome Resources. Compared to fungi and animal mitochondrial genomes, plant mitochondrial genomes are more complex and encode a higher number of genes. These genomes play vital roles in plant development and productivity [1–3]. There is an unusual size variation found in seed plant mt genomes, especially within the same family [4]. Seed plant mt genomes are distinctive for their frequent insertion of foreign DNA by gene transfer [5], very low mutation...
rate [6] and dynamic structure [7]. Terrestrial seed plants have increased their mt genome sizes by adopting new mechanisms to facilitate more gene exchange between nuclear genomes and mt genomes as well as cp genomes and mt genomes [8]. Investigations of the mt genomes of different important angiosperm species, including *Beta vulgaris* [9], *Arabidopsis thaliana* [10], *Brassica napus* [11, 12], *Triticum aestivum* [1], *Oryza sativa* [13, 14], *Zea mays* [15–17], *Vitis vinifera* [18], *Nicotiana tabacum* [19], *Vigna radiata* [20] and *Citrullus lanatus* [4], have been performed. Together with physical mapping [21-23], these mt genomes showed various properties, such as a slow rate of evolution, large genome size (200–2400 kb), multipartite structure, uptake of foreign DNA and different modes of gene expression (RNA editing and splicing) [24]. However, the above properties are unable to explain the diversity of mitotypes within each plant genus and species.

Much of the structure and size difference in plant mt genomes are repeated sequences in the DNA content [25]. The repeated DNA sequences are also sources for intragenomic recombination, and they trigger various changes in mitochondrial genome evolution and structural dynamism [26]. To analyze the evolutionary distinctiveness of a plant’s mitochondrial genome within one plant genus or species, more systematic and specific sequences are needed. To date, there are no specific and organized sequences for an angiosperm genus to analyze multiple species for mitochondrial genome derivation; therefore, the mechanism of having such uniqueness has not yet been revealed [26].

Previously, phylogenetic analysis [27-29] has reported the complicated and laborious method of amplifying selected loci in genes, some of which are unable to provide sufficient and accurate information about phylogenetic resolutions. Recently, next generation sequencing advancements have led to various organelle genome sequencing, which is continuously contributing to various areas of biology. The use of whole organelle genome sequencing, especially chloroplast and mitochondria genomes, has been recently demonstrated as a potential barcode [30] that can assist in overcoming the previous process of collecting data over generations. Furthermore, due to recombination in the nucleus, data may lead to unreliable phylogenies; organelles are structurally stable, non-recombinant, haploid and have certain advantages in phylogenetic reconstructions [31].

Rice is one of the most important cereal crops, a staple food for more than half of the world’s population and a model crop for cereal genomics. The genus *Oryza* has two cultivated species and more than 20 wild relatives based on pairs of chromosomes. *Oryza* species are categorized into 10 genome types: AA, BB, CC, BBCC, EE, FF, GG, CCDD, HHJJ and HHKK [27]. Furthermore, these genome types have different species and subspecies. *O. sativa*, one of the important species that has an AA genome type, is further divided into the following subspecies: *O. sativa* L. spp. *japonica* and *O. sativa* L. spp. *indica*, which has a global distribution [32]. Moreover, wild *Oryza* species have the potential to resist against biotic and abiotic stresses, especially to insect pests (Heinrichs et al., 1985). *O. minuta*, a tetraploid wild relative of cultivated rice with a BBCC genome, exhibits the potential to resist against blast blight, bacterial blight (BB), white backed planthopper (WBPH) and brown plant hopper (BPH) diseases. Furthermore, various resistance genes have been transferred successfully to cultivated rice from *O. minuta* [33, 34]. These wild and cultivated species share a valuable genetic diversity that has contributed greatly to the improvement of rice crops. To identify more desired genes and ensure effective conservation, analysis of their phylogenetic and evolutionary relationship is very important [35]. Hence, the current study was performed. Various organelle genomes of Poaceae members have already been reported, including *O. sativa indica*, *O. rufipogon*, *O. sativa japonica*, *T. aestivum* and *Z. mays* [13, 36–38]. Recently, many nuclear genomes from various economically important plants have been published or are still in progress [39]. Billions of short read sequencing data for the whole genome from many species are deposited in a
public database. In this study, we aimed to analyze the complete mitochondrial genome sequence of *O. minuta* (mtDNA) and compare it with other sequenced mt genomes of the Poaceae family. The current study will provide information for the further understanding of mt genome evolution in related species.

**Materials and Methods**

In this study, we successfully assembled and annotated the complete mtDNA of a wild cultivar of *O. sativa* (IRGC 101140) following the method described previously [40, 41]. Approximately sixty million raw Illumina reads were demultiplexed and trimmed. The raw reads were filtered and then assembled *de novo* into contigs using CLC Genomics Workbench v7.0 (CLC Bio, Aarhus, Denmark).

**Sequence data analysis**

BLAST searches were conducted on all of the contigs using the NCBI database (http://www.ncbi.nlm.nih.gov/) for the annotation of mitochondrial sequences using previous angiosperm annotated mitochondrial genes as query sequences. tRNA scan-SE software (http://lowelab.ucsc.edu/trNAscan-SE/) was used to identify tRNAs in the genome. The ORF-Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) was used to predict and annotate open reading frames (ORFs) with a minimum size of 100 codons. Analysis of repeat sequences was performed as described previously [42]. While the circular map of mt genome was created using OGDraw v1.2 (http://ogdraw.mpimp-golm.mpg.de/), the tandem repeats were identified with Tandem Repeat Finder (TRF) using a default setting [43]. The tandem repeat lengths were set to 20 bp or more with a maximum period size and a minimum alignment score of 500 and 50, respectively, and the repeats identity was set to >80%. The annotated genome sequence was submitted to NCBI with the GenBank accession No. KU176938.

**Comparing mitochondrial genomes and evolutionary analysis**

The *O. minuta* mitochondrial genome (GenBank: KU176938) sequence described here was compared with seven other reported Poaceae mitotypes: *O. sativa japonica* (GenBank: BA000029), *O. sativa indica* (GenBank: DQ167399), *O. rufipogon* (GenBank: AP011076), *Triticum aestivum* (GenBank: NC007579), *Zea mays* spp. parviglumis (GenBank: NC008332), *Z. mays* spp. *mays* (GenBank: NC007982) and *Sorghum bicolor* (GenBank: NC008360), using NCBI-blastn. For comparison, 20 protein coding genes (*atp9*, *ccmC*, *ccmFN1*, *cox1*, *cox2*, *cox3*, *cob*, *matR*, *nad4L*, *nad6*, *nad9*, *rps1*, *rps3*, *rps7*, *rps12*, *rps13*, *rps4*, *rnr5*, *rpl2*, and *rpl5*), which were shared by these eight species, were extracted and successively joined together. MEGA 6 [44] was used to construct a neighbor-joining tree [45] with 1000 bootstrap replications [46]. For the whole genome as well as the coding regions, comparison distance matrices were computed using Progressive Mauve (The Darling lab at the University of Technology Sydney), and then the whole genome distance matrix was converted to a heat map [47]. Furthermore, the dot matrix method was also used to analyze similarities among different sequences [48].

**Results**

**Mitochondrial genome of *O. minuta***

Mitochondrial DNA of *O. minuta* was assembled into a circular genome of 515,022 bp with 44% overall GC content, which is almost similar to the mtDNA of other related species (Table 1). The non-coding sequences of *O. minuta* mtDNA is almost 86.04%, which is less
than the previously reported angiosperm average non-coding sequences content (89.46%) [29]. Genes account for 13.9% of the genome and 71,846 bp in length.

Gene content and open reading frames (ORFs)
A total of 93 genes were identified, including 60 protein-coding genes (PCGs), 31 tRNA genes and 2 rRNA genes using BLAST and TRNA scan-SE (Fig 1 and Table 1). The protein coding genes were present in a range of 225 bp (atp9) to 8,980 bp (nad4), which included 31 genes for the production of ATP synthase and the electron transport chain, consisting of the following subunits: 4 subunits of complex I (nad3, 4, 6, 9), 1 subunit of complex III (cob), 3 subunits of complex IV (cox1-3) and 1 subunit of complex V (atp9) (Table 2). Furthermore, there were four genes for cytochrome c biogenesis (ccmB, ccmC, ccmFN and ccmFC), three genes for large ribosomal proteins (rpl2, 5, 16) and seven genes for small ribosomal protein (rps1, 3, 4, 7, 12, 13, and 19) (Table 2). The total length of the 60 protein coding genes of O. minuta mtDNA was 71,846 bp (Table 1), accounting for 13.9% of its total genome length, which is different from other Oryza genus mitogenomes. In O. minuta, the most common start codon for the protein coding genes was ATG, except for ccmB (start codon ATC), matR (start codon AGA) and rrn5 (start codon AAA), as reported previously (Handa, 2003). Ten genes (ccmB, cox3/3, orf160, orf194, orf241, rps1, rps12/12 and rps13) were expected to terminate with TGA and eleven (ccmC, ccmFn, cob, cox1, cox2, mat-R orf25/orf153, orf194, orfx, rps3) with TAG; other protein coding genes use TAA as their termination codon.

A total of 31 tRNA sequences (2,328 bp) were found in O. minuta mtDNA (Table 1) in the range of 71–88 bp in length. The GC content of the tRNA genes was 51.3%, with A, C, G, and T as 22.1, 22.6, 28.7 and 26.6%, respectively, which was higher than the overall GC composition of the mtDNA. Among these genes, tRNAs for 16 amino acids, including seven for Methionine (Met), three for serine (Ser), two for Lysine (Lys) and Cysteine (Cys), are encoded (Table 3). The genome deficient tRNAs for the following amino acids were: Valine (Val), Alanine (Ala), Leucine (Leu), Glycine (Gly), Tryptophan (Trp) and Threonine (Thr) (Table 3). A total of 627 ORFs were identified, which was longer than 100 codons in the O. minuta mitochondrial genome using ORF-Finder. All ORFs were a single copy between 200 and 800 bp in length, except for nine ORFs that were longer than 2,000 bp, including the 2,508 bp orf492 and the 3,474 bp orf5.

Repétitive sequences in the mitochondrial genome
Large repeats were identified in the mitochondrial genome of O. minuta. Seven pairs of repetitive sequences were found and designated as R1-R7 (Table 4). R1-R4 (19,773, 16,451, 7,984, 6,856 bp) had pairs of large repeats in the same direction longer than 6,050 bp, while R5-R7

| Features                  | O. minuta | O. sativa j | O. sativa i | O. rufipogon |
|---------------------------|-----------|-------------|-------------|--------------|
| Genome size (bp)          | 515,022   | 490,520     | 491,515     | 559,045      |
| GC contents               | 44        | 43.9        | 43.8        | 44           |
| Total gene contents       | 93        | 81          | 94          | 59           |
| Protein coding gene       | 60        | 56          | 53          | 41           |
| rRNA                      | 2         | 3           | 6           | 3            |
| tRNA                      | 31        | 22          | 33          | 15           |
| Total gene length         | 71,846    | 53,182      | 156,514     | 43,715       |

doi:10.1371/journal.pone.0152937.t001
Fig 1. Mitochondria genome map of *O. minuta*. Features on the clockwise- and counter-clockwise transcribed strands are drawn on the inside and outside of the circle, respectively.

doi:10.1371/journal.pone.0152937.g001
had inverted repeats longer than 70 bp. The large repeat from R2 contained two genes, \textit{trnE} and \textit{orf173}, while R3 contained the \textit{trnK} gene. No protein-coding gene was found in the other repeats. Furthermore, a total of 22 tandem repeats of more than 10 bp were identified in the \textit{O. minuta} mitogenome (Table 5). The length of the repeat units in these regions varied between 11 and 70 bp, and up to 3 repeat units presented with having a varied identity percentage from 80 to 100% (Table 5).

**\textbf{O. minuta} mtDNA comparison with other \textit{Oryza} species**

We compared the sequences of the mtDNA from \textit{O. minuta} (515,022 bp) with three \textit{Oryza} species: \textit{O. sativa indica}, \textit{O. sativa japonica} and \textit{O. rufipogon}, which all have circular mitochondrial genomes of 491,515, 490,520 and 559,045 bp, respectively (Table 1). The mitochondrial genome of \textit{O. minuta} was larger than \textit{O. sativa indica} and \textit{O. sativa japonica}, while smaller than \textit{O. rufipogon} (Table 1). The GC content of \textit{O. minuta} was slightly different from other mitogenomes. Similarly, nucleotide base content of the total length of the genes with known functions (71,846 bp) was different among these mitogenomes (Table 1). Analysis of the genes with known functions showed that \textit{O. minuta} shared 24 protein encoding genes, 15 tRNA genes and 1 ribosome gene (Table 2); paralogous genes that presented in more than one copy were counted here as one gene. The numbers of genes with known functions were almost the same in these mitogenomes, but the total number of genes varied, ranging from 59 in \textit{O. rufipogon} to 94 in \textit{O. sativa indica} (Table 1).
Evolutionary relationships of the *O. minuta* mitogenome

To explain the evolutionary relationship of *O. minuta* within the Poaceae family, the mtDNA genomes of selected species were compared with related mtDNA sequences using blastn. Similar regions in these mitogenomes were aligned to the mtDNA of *O. minuta* (reference genome).

Table 3. Recognition of anticodons by tRNA genes found in the mitochondrial genome of *O. minuta*.

| NAME | Type | Anticodon | Length(bp) | Orientation |
|------|------|-----------|------------|-------------|
| trnQ | Gln  | (TTG)     | 72         | Inverted    |
| trnH | His  | (GTG)     | 74         | Inverted    |
| trnM | Met  | (CAT)     | 73         | Direct      |
| trnP | Pro  | (TGG)     | 75         | Direct      |
| trnS | Ser  | (GCT)     | 88         | Inverted    |
| trnM) | Met  | (CAT)     | 74         | Direct      |
| trnS | Ser  | (GGA)     | 87         | Inverted    |
| trnF | Phe  | (GAA)     | 73         | Inverted    |
| trnY | Tyr  | (GTA)     | 83         | Inverted    |
| trnE | Glu  | (TTC)     | 72         | Inverted    |
| trnC | Cys  | (GCA)     | 71         | Direct      |
| trnR | Arg  | (TCT)     | 72         | Direct      |
| trnS | Ser  | (GGA)     | 87         | Inverted    |
| trnF | Phe  | (GAA)     | 73         | Inverted    |
| trnY | Tyr  | (GTA)     | 83         | Inverted    |
| trnE | Glu  | (TTC)     | 72         | Inverted    |
| trnC | Cys  | (GCA)     | 71         | Direct      |
| trnM | Met  | (CAT)     | 74         | Inverted    |
| trnK | Lys  | (TTT)     | 73         | Inverted    |
| trnN | Asn  | (GTT)     | 72         | Inverted    |
| trnD | Asp  | (GTC)     | 74         | Direct      |
| trnI | Ile  | (CAT)     | 74         | Inverted    |
| trnQ | Gln  | (TTG)     | 72         | Inverted    |
| trnH | His  | (GTG)     | 74         | Inverted    |
| trnM | Met  | (CAT)     | 73         | Direct      |
| trnP | Pro  | (TGG)     | 75         | Direct      |
| trnM | Met  | (CAT)     | 74         | Inverted    |
| trnK | Lys  | (TTT)     | 73         | Inverted    |
| trnN | Asn  | (GTT)     | 72         | Inverted    |
| trnD | Asp  | (GTC)     | 74         | Direct      |
| trnI | Ile  | (CAT)     | 74         | Inverted    |

doi:10.1371/journal.pone.0152937.t003

Table 4. Large repeats in the mitochondrial genome of *O. minuta*.

| No  | Type | Size  | Copy-1                        | Copy-2                        | Difference between copies | Identity |
|-----|------|-------|-------------------------------|-------------------------------|---------------------------|----------|
| R1  | DR   | 19773 | 20563–40336                   | 189513–209286                 | identical                 | 100      |
| R2  | DR   | 16451 | 279979–296430                 | 343373–359824                 | identical                 | 100      |
| R3  | DR   | 7984  | 3630–11614                    | 172517–180501                 | identical                 | 100      |
| R4  | DR   | 6856  | 55990–62846                   | 225087–231943                 | identical                 | 100      |
| R5  | IR   | 112   | 408736–408848                 | 235292–235374                 | identical                 | 100      |
| R6  | IR   | 82    | 57033–57124                   | 66037–66126                   | 2bp                       | 99%      |
| R7  | IR   | 70    | 47047–47116                   | 51677–51746                   | identical                 | 100%     |

doi:10.1371/journal.pone.0152937.t004
The *O. minuta* sequence showed 89% identity to that of the *O. rufipogon* mtDNA sequences. The sequence identity shared by the mtDNA of *O. minuta* with *O. sativa* (indica and japonica), *S. bicolor*, *T. aestivum* and *Z. mays* (mays and parviglumis) were 34.2, 34.2, 10.1, 17.8, 14.1 and 13.3%, respectively (S1 Fig). These results strongly suggested that *O. minuta* was closely related to *O. rufipogon*, and the evolutionary relationship between these two was much stronger than that of *O. minuta* with any other species.

To support these results, a dot matrix analysis showed that the length of syntenic regions of *O. minuta* with the *O. rufipogon* mitogenome were longer and straight. Additionally, the identity of *O. minuta* with *O. sativa indica* and *japonica* was lower, and the distribution of the syntenic regions was more dispersed than that of *O. rufipogon* (Fig 2A–2C). Moreover, the phylogenetic relationships among the Poaceae members (Fig 3; S2 Fig) were conducted using 20 conserved genes among the reported mitogenomes by the neighbor-joining method. These results were consistent with our comparative results based on mitochondrial genome analysis and revealed that *O. minuta* was more closely related to *O. rufipogon* than any other Poaceae member.

**Discussion**

The Poaceae is an important plant family with significant importance to human beings because the plants in this family are the main sources for human food and animal feed. The rapid increases in genomic analysis and bioinformatics approaches have revealed the great agronomic and scientific importance of this model species. Furthermore, mitochondrial genome analysis of closely related species has significantly improved the knowledge of molecular evolution and phylogenetic analysis [49] in various species. *O. minuta*, a wild rice with the BBCC
Mitochondrial Genome Analysis of Wild Rice

The mitochondrial genomes of Poaceae members were generally larger in size (452.52–704 kb) compared with other angiosperm plants. The *O. minuta* mtDNA (515 kb) was larger than *O. sativa* (*indica and japonica*) while smaller than *O. rufipogon* (559.04 kb). Similarly, the percentage of the GC contents were slightly different from *O. sativa* (*indica and japonica*) (43.8

**Fig 2.** Dot matrix alignment of the *O. minuta* (x-axis) with other mitochondrial genomes of Poaceae members (y-axis). (A) *O. rufipogon*, (B) *O. sativa japonica*, (C) *O. sativa indica*, (D) *S. bicolor*, (E) *T. aestivum*, (F) *Z. mays* spp. *parviglumis* and (G) *Z. mays* spp. *mays*.

doi:10.1371/journal.pone.0152937.g002
and 43.9) and similar with *O. rufipogon* [8, 13, 36]. Comparison of the *O. minuta* mtDNA with the three mitogenomes above revealed that the protein coding genes were relatively conserved among these mitogenomes. A total of 24 coding genes, 15 tRNA genes and 1 ribosomal rRNA gene were shared within these mitogenomes. However, various genes (*ccmB*, *ccmC*, *cox2*, *cox3*, *nad3*, *nad4*, *rpl16*, *rps12*, *rps3*, *rps7* and *rrn5*) were present in the duplication of the *O. minuta* genome. Furthermore, genes (*ccmB*, *ccFc*, *nad4*, and *rpl16*) were absent in the *O. rufipogon* mitogenome [8] (Table 2).

A total of 31 tRNA sequences for 15 genes were identified in *O. minuta* mtDNA, accounting for only 0.40% of the mitochondrial genome (Table 1). Among them, six genes exhibited high sequence similarity (>99%) to the chloroplast genome and seemed to be derived from the chloroplast. The chloroplast-derived genes (*trnW-CCA*, *trnD-GUC*, *trnS-GGA*, *trnH-GTG*, *trnM-CAT* and *trnL-CAA*), which are commonly found in angiosperm mitochondrial genomes [51], were present in the *O. minuta* mitogenome. Furthermore, another two genes, *trnQ-UUG*...
and trnP-GGG reported in dicot transfer events [52, 53], were additionally found in O. minuta. Thus, functional tRNA for eight amino acids (trnB, trnA, trnT, trnV, trnZ, trnW and trnX) were absent from the O. minuta mitogenome, although for protein synthesis in the mitochondria, tRNAs for 20 amino acids are necessary. These results revealed that the nuclear genome might have supplied these missing tRNAs. Thus, nine tRNAs involved in mitochondria biogenesis in rice are of mitochondrial origin, six are of plastid origin and the above missing are probably of nuclear origin. These results paralleled results previously reported for the O. sativa mitogenome [13]. Furthermore, previously reported trnS and trnM for rice mitochondria and plastid like tRNAs, respectively [54], were additionally identified in the O. minuta mitogenome.

Searching for repeated sequences showed four direct and three inverted repeats longer than 6,050 and 70 bp, respectively (Table 4). The longest inverted and direct repeats that showed 100% identity were 112 and 19,773 bp long, respectively. Similarly, a total of 22 tandem repeats longer than 10 bp were additionally identified in the O. minuta mitogenome (Tables 4 and 5). These results were different from those previously reported for the O. sativa mitogenome, which had direct and inverted repeats of 45,584 and 946 bp, respectively [13]. Furthermore, the multipartite structure of the plant mitochondrial genome is thought to be generated through the recombination of repeated sequences; however, the involvement of these sequences in rice mtDNA is not yet clear [13]. Furthermore, the phylogenetic analysis of O. minuta’s complete mtDNA as well as 20 conserved genes with other related species revealed that it was closer to O. rufipogon than to any other related species.

**Conclusion**

In this study, we reported the complete mitochondrial genome of O. minuta. The O. minuta mtDNA is composed of 515,022 bp and contained 60 known protein coding genes, two rRNA (5rRNA) and 31 tRNA genes. Genome organization and gene content is typical of the Oryza species and highly similar to that of O. rufipogon (89% identical at the nucleotide level). Furthermore, it shared 24 protein-coding genes, 15 tRNA genes and 1 ribosomal RNA gene with other O. sativa (indica and japonica). Similarly, the evolutionary relationship analysis with other Poaceae members revealed that the mtDNA of O. minuta is closely related to O. rufipogon. This study will improve our understanding of O. minuta (wild rice) and the evolution of the mitogenomes within the Poaceae family.

**Supporting Information**

S1 Fig. Heat map based on a pair-wise distance matrix of whole mitogenomes alignment as computed by Progressive Mauve. Mitochondrial genome alignments were performed using O. minuta as a reference genome for the other seven Poaceae members. Distance values correspond to a gradient of color steps ranging from light gray (lowest distance) to dark black (highest distance value).

(TIF)

S2 Fig. Comparison of O. minuta mitogenome coding regions with other mitogenomes of Poaceae members using Progressive Mauve alignments.

(TIF)

**Author Contributions**

Conceived and designed the experiments: SA IJL. Performed the experiments: SA ALK ARK MW SMK MAK RS CWS JHS IJL. Analyzed the data: SA. Contributed reagents/materials/analysis tools: IJL. Wrote the paper: SA MW ALK.
References

1. Ogihara Y, Yamazaki Y, Murai K, Kanno A, Terachi T, Shiina T, et al. Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. Nucleic Acids Res. 2005; 33(19):6235–50. Epub 2005/11/02. doi:10.1093/nar/gki925 PMID:16260473; PubMed Central PMCID: PMCPMC1275586.

2. Li L, Wang B, Liu Y, Qiu YL. The complete mitochondrial genome sequence of the hornwort Megaceros aenigmaticus shows a mixed mode of conservative yet dynamic evolution in early land plant mitochondrial genomes. J Mol Evol. 2009; 68(6):665–78. Epub 2009/05/29. doi:10.1007/s00239-009-9240-7 PMID:19475442.

3. Liu Y, Xue JY, Wang B, Li LB, Qiu YL. The Mitochondrial Genomes of the Early Land Plants Treubia lacunosa and Anomodon rugelii: Dynamic and Conservative Evolution. Plos One. 2011; 6(10). ARTN e25836 doi:10.1371/journal.pone.0025836. WOS:000295966900062.

4. Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD. Insights into the Evolution of Mitochondrial Genome Size from Complete Sequences of Citrullus lanatus and Cucurbita pepo (Cucurbitaceae). Mol Biol Evol. 2010; 27(6):1436–48. doi:10.1093/molbev/msq029 PMID:20118192

5. Goremykin VV, Salamini F, Velasco R, Viola R. Mitochondrial DNA of Vitis vinifera and the Issue of Rampant Horizontal Gene Transfer. Mol Biol Evol. 2009; 26(1):99–110. doi:10.1093/molbev/msn226 PMID:18922764

6. Palmer J, Herbon L. Plant mitochondrial DNA evolved rapidly in structure, but slowly in sequence. J Mol Evol. 1988; 28(1–2):87–97. doi:10.1007/BF02143500 PMID:3148746

7. Matter F. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences. Preface. Philos Trans R Soc Lond B Biol Sci. 1988; 319(83–84).

8. Stern DB, Palmer JD. Tripartite mitochondrial genome of spinach: physical structure, mitochondrial gene mapping, and locations of transposed chloroplast DNA sequences. Nucleic Acids Res. 1986; 14(14):5651–66. Epub 1986/07/25. PMID:3016660; PubMed Central PMCID: PMCPMC311583.

9. Kubo T, Nishizawa S, Sugawara A, Itchoda N, Estiati A, Mikami T. The complete nucleotide sequence of the mitochondrial genome of sugar beet (Beta vulgaris L.) reveals a novel gene for tRNA(Cys)(GCA). Nucleic Acids Res. 2000; 28(13):2571–6. Epub 2000/06/28. PMID:10871408; PubMed Central PMCID: PMCPMC102699.

10. Unseld M, Marienfeld JR, Brandt P, Brennicke A. The mitochondrial genome of Arabidopsis thaliana contains 57 genes in 366,924 nucleotides. Nat Genet. 1997; 15(1):57–61. PMID:988169

11. Handa H. The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (Brassica napus L.): comparative analysis of the mitochondrial genomes of rapeseed and Arabidopsis thaliana. Nucleic Acids Res. 2003; 31(20):5907–16. Epub 2003/10/08. PMID:14530439; PubMed Central PMCID: PMCPMC219474.

12. Chen JM, Guan RZ, Chang SX, Du TO, Zhang HS, Xing H. Substoichiometrically Different Mitotypes Coexist in Mitochondrial Genomes of Brassica napus L. Plos One. 2011; 6(3). ARTN e17662 doi:10.1371/journal.pone.0017662. WOS:000288219100022.

13. Notsu Y, Masood S, Nishikawa A, Itchoda N, Estiati A, Mikami T, et al. The complete nucleotide sequence of the mitochondrial genome of sugar beet (Beta vulgaris L.) reveals a novel gene for tRNA(Cys)(GCA). Nucleic Acids Res. 2000; 28(13):2571–6. Epub 2000/06/28. PMID:10871408; PubMed Central PMCID: PMCPMC102699.

14. Fujii S, Kazama T, Yamada M, Toriyama K. Discovery of global genomic re-organization based on comparison of two newly sequenced rice mitochondrial genomes with cytoplasmic male sterility-related genes. Bmc Genomics. 2010; 11:209-. doi:10.1186/1471-2164-11-209. PMC2851602. PMID:20346185

15. Allen JO, Fauron CM, Minx P, Roark L, Oddiraju S, Lin GN, et al. Comparisons Among Two Fertile and Three Male-Sterile Mitochondrial Genomes of Maize. Genetics. 2007; 177(2):1173–92. doi:10.1534/genetics.107.07312. PMC2034622. PMID:17660568

16. Darraçq A, Varré J-S, Touzet P. A scenario of mitochondrial genome evolution in maize based on rearrangement events. Bmc Genomics. 2010; 11:233-. doi:10.1186/1471-2164-11-233. PMID:20380689

17. Clifton SW, Minx P, Fauron CMR, Gibson M, Allen JO, Sun H, et al. Sequence and Comparative Analysis of the Maize NB Mitochondrial Genome. Plant Physiol. 2004; 136(3):3486–503. doi:10.1104/pp.104.044602. PMC527149. PMID:15542500

18. Schnell J, Steele M, Bean J, Neuspiel M, Girard C, Dormann N, et al. A comparative analysis of insertional effects in genetically engineered plants: considerations for pre-market assessments. Transgenic
Sugiyama Y, Watase Y, Nagase M, Makita N, Yagura S, Hirai A, et al. The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. Mol Genet Genomics. 2005; 272(6):603–15. Epub 2004/12/08. doi: 10.1007/s00438-004-1075-8 PMID: 15583938.

Alverson AJ, Zhuo S, Rice DW, Sloan DB, Palmer JD. The Mitochondrial Genome of the Legume Vigna radiata and the Analysis of Recombination across Short Mitochondrial Repeats. Plos One. 2011; 6(1):e16404. doi: 10.1371/journal.pone.0016404. PMC3024419. PMID: 21283772.

Palmer JD, Shields CR. Tripartite structure of the Brassica campestris mitochondrial genome. Nature. 1984; 307(5950):437–40.

Chetritl P, Mathieu C, Muller JP, Vedel F. Physical and gene mapping of cauliflower (Brassica oleracea) mitochondrial DNA. Curr Genet. 1984; 8(6):413–21. Epub 1984/08/01. doi: 10.1007/bf00433907 PMID: 24177911.

Palmer JD, Herbon LA. Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J Mol Evol. 1988; 28(1–2):87–97. Epub 1988/12/01. PMID: 3147846.

Schuster W, Brennicke A. The plant mitochondrial genome—physical structure, information-content, mRNA editing, and gene migration to the nucleus. Annual Review of Plant Physiology and Plant Molecular Biology. 1994; 45:61–78. WOS:A1994NU95900004.

Palmer JD, Herbon LA. Unicircular structure of the Brassica hirta mitochondrial genome. Curr Genet. 1987; 11(6–7):565–70. Epub 1987/01/01. PMID: 2836082.

Ward BL, Anderson RS, Bendich AJ. The mitochondrial genome is large and variable in a family of plants (cucurbitaceae). Cell. 1981; 25(3):793–803. Epub 1981/09/01. PMID: 6269758.

Ge S, Sang T, Lu BR, Hong DY. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proc Natl Acad Sci U S A. 1999; 96(25):14400–5. Epub 1999/12/10. PMID: 10588717; PubMed Central PMCID: PMC244448.

Zhu T, Xu PZ, Liu JP, Peng S, Mo XC, Gao LZ. Phylogenetic relationships and genome divergence among the AA-genome species of the genus Oryza as revealed by 53 nuclear genes and 16 intergenic regions. Mol Phylogenet Evol. 2014; 70:348–61. Epub 2013/10/24. doi: 10.1016/j.ympev.2013.10.008 PMID: 24148990.

Zhu Q, Ge S. Phylogenetic relationships among A-genome species of the genus Oryza revealed by intron sequences of four nuclear genes. New Phytol. 2005; 167(1):249–65. Epub 2005/06/14. doi: 10.1111/j.1469-8137.2005.01406.x PMID: 15948847.

Nock CJ, Waters DLE, Edwards MA, Bowen SG, Rice N, Cordeiro GM, et al. Chloroplast genome sequences from total DNA for plant identification. Plant Biotechnol J. 2011; 9(3):328–33. doi: 10.1111/j.1467-7652.2010.00558.x. WOS:000288630900004. PMID: 20796245.

Small RL, Cronn RC, Wendel JF. Use of nuclear genes for phylogeny reconstruction in plants. Australian Systematic Botany. 2004; 17(2):145–70. doi: 10.1071/SB03015

Chang T-T. The origin, evolution, cultivation, dissemination, and diversification of Asian and African rice. Euphytica. 1976; 25(3):425–41. doi: 10.1007/BF00041576.

Amante-Bordeos A, Sitch LA, Nelson R, Dalmacio RD, Oliva NP, Aswidinnoor H, et al. Transfer of bacterial blight and blast resistance from the tetraploid wild rice Oryza minuta to cultivated rice, Oryza sativa. Theor Appl Genet. 1992; 84(3–4):345–54. Epub 1992/07/01. doi: 10.1007/bf00229493 PMID: 24203194.

Rahman ML, Jiang W, Chu SH, Qiao Y, Ham TH, Woo MO, et al. High-resolution mapping of two rice brown planthopper resistance genes, Bph20(l) and Bph21(l), originating from Oryza minuta. Theor Appl Genet. 2009; 119(7):1237–46. Epub 2009/08/12. doi: 10.1007/s00122-009-1125-z PMID: 19669727.

Guo SB, Wei Y, Li XQ, Liu KQ, Huang FK, Chen CH, Gao GQ. Development and Identification of Introgression Lines from Cross of Oryza sativa and Oryza minuta. Rice Science. 2013; 20(2):95–102. doi: 10.1007/s11726-013-6011-0

Tian X, Zheng J, Hu S, Yu J. The Rice Mitochondrial Genomes and Their Variations. Plant Physiol. 2006; 140(2):401–10. doi: 10.1104/pp.105.070600. PMC1363132. PMID: 16384910.

Cui P, Liu H, Lin Q, Ding F, Zhuo S, Hu S, et al. A complete mitochondrial genome of wheat (Triticum aestivum cv. Chinese Yumai), and fast evolving mitochondrial genes in higher plants. J Genet. 2009; 88(3):299–307. Epub 2010/01/21. PMID: 20086295.

Allen JO, Fauron CM, Minx P, Roark L, Oddiraju S, Lin GN, et al. Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. Genetics. 2007; 177(2):1173–92. doi: 10.1534/genetics.107.073312. WOS:000250657800043. PMID: 17660568.

Michael TP, Jackson S. The First 50 Plant Genomes. The Plant Genome. 2013; 6(2).
40. Wu Z. The whole chloroplast genome of shrub willows (Salix suchowensis). Mitochondr DNA. 2014:1–2. Epub 2014/11/25. PMID: 25418623.
41. Wu Z. The completed eight chloroplast genomes of tomato from Solanum genus. Mitochondr DNA. 2015:1–3. Epub 2015/01/22. doi: 10.3109/19401736.2014.1003890 PMID: 25604480.
42. Chang S, Yang T, Du T, Huang Y, Chen J, Yan J, et al. Mitochondrial genome sequencing helps show the evolutionary mechanism of mitochondrial genome formation in Brassica. BMC Genomics. 2011; 12:497. Epub 2011/10/13. doi: 10.1186/1471-2164-12-497 PMID: 21988783; PubMed Central PMCID: PMC3204307.
43. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999; 27 (2):573–80. Epub 1998/12/24. PMID: 9862982; PubMed Central PMCID: PMCPMC148217.
44. Kumar S, Nei M, Dudley J, Tamura K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform. 2008; 9(4):299–306. Epub 2008/04/18. doi: 10.1093/bib/bbn017 PMID: 18417537; PubMed Central PMCID: PMCPMC2562624.
45. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987; 4(4):406–25. Epub 1987/07/01. PMID: 21988783.
46. Felsenstein J. Confidence-Limits on Phylogenies—an approach using the bootstrap. evolution 39: 783–791. doi:10.2307/2408678 View Article PubMed. NCBI Google Scholar. 1985.
47. Darling AE, Mau B, Perna NT. progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. PloS One. 2010; 5(6). ARTN e11147 doi:10.1371/journal.pone.0011147. WOS:000279140800001.
48. Noe L, Kucherov G. YASS: enhancing the sensitivity of DNA similarity search. Nucleic Acids Res. 2005; 33:W540–W3. doi: 10.1093/nar/gki478. WOS:000230271400110. PMID: 15980530.
49. Yuan ML, Wei DD, Wang BJ, Dou W, Wang JJ. The complete mitochondrial genome of the citrus red mite Panonychus citri (Acari: Tetranychidae): high genome rearrangement and extremely truncated tRNAs. BMC Genomics. 2010; 11. Art 597 doi: 10.1186/1471-2164-11-597. WOS:000284117600001.
50. Vaughan DA. The wild relatives of rice: a genetic resources handbook. Manila: International Rice Research Institute; 1994. vii + 137 pp. p.
51. Chang S, Wang Y, Lu J, Gai J, Li J, Chu P, et al. Correction: The Mitochondrial Genome of Soybean Reveals Complex Genome Structures and Gene Evolution at Intercellular and Phylogenetic Levels. Plos One. 2015; 8(6).
52. Zhang T, Fang Y, Wang X, Deng X, Zhang X, et al. (2012) The complete chloroplast and mitochondrial genome sequences of Boea hygrometrica: insights into the evolution of plant organelar genomes. PLoS One 7:e30531–e30531. doi: 10.1371/journal.pone.0030531 PMID: 22291979.
53. Goremykin VV, Lockhart PJ, Viola R, Velasco R. The mitochondrial genome of Malus domestica and the import-driven hypothesis of mitochondrial genome expansion in seed plants. The Plant Journal. 2012; 71(4):615–26. doi: 10.1111/j.1365-313X.2012.05014.x PMID: 22469901.
54. Miyata S, Nakazono M, Hirai A. Transcription of plastid-derived tRNA genes in rice mitochondria. Curr Genet. 1998; 34(3):216–20. Epub 1998/09/24. PMID: 9745025.