

Limited Polymorphisms between Two Whole Plastid Genomes in the Genus *Zizania* (Zizaniinae)

Zhiqiang Wu^{1,3*#}, Cuihua Gu^{2,3*#}, Luke R Tembrock³ and Song Ge¹

¹State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China

²School of Landscape and Architecture, Zhejiang Agriculture and Forestry University, Hangzhou 311300, P.R. China.

³Department of Biology, Colorado State University, Fort Collins, Colorado, USA

Contributed equally to this work

Abstract

Plastids, originated from cyanobacteria through endosymbiosis, contain their own DNA genome (plastome), and are uniparentally inherited in most of plant species. In this study, we reported the complete plastome sequence of Manchurian wild rice (*Zizania latifolia*) obtained by traditional Sanger sequencing and compared it with the previously published plastome of North America wild rice (*Z. aquatica*) from the same genus. The plastome of *Z. latifolia* has a total sequence length of 136,461 bp exhibiting a typical circular structure including a pair of 20,878 bp inverted repeats (IRa, b) separated by a large single-copy region (LSC) of 82,115 bp and a small single-copy region (SSC) of 12,590 bp, and it is only 97 bp longer than North America wild rice. The gene content, order, and orientation are similar to all other grass species. Four junctions in two plastome structures are exactly the same between the *Zizania* species. From complete genome comparisons, 744 (568, 46 and 130 for LSC, IR and SSC) substitutions are found between the two *Zizania* plastomes including 267 from coding regions and 477 from non-coding regions. Insertions/deletions (Indels) are mainly found in non-coding regions (136) except the only one repeat indel from coding regions in *rpoC2* gene. The most informative biomarkers (*rps16-trnQ*, *ndhF* and *matK*) are also the most divergent between the two *Zizania* species. The completed plastome and comparative analyses from these two *Zizania* species would be as important resource for further systematic or population genetic studies in the Zizaniinae subtribe, marker assisted breeding, and plastomic transformation.

Keywords: Oryzeae; *Zizania*; PCR; Plastome; Polymorphisms; Indels

Introduction

The tribe Oryzeae contains some of the most economically important species in the world. The tribe is divided into two well supported subtribes named Oryzinae and Zizaniinae [1-4]. Economically important species are found in both subtribes such as cultivated rice in Asian and African (*Oryza sativa* and *O. glaberrima*) from Oryzinae and the wild rice (*Zizania latifolia* and *Z. aquatica*) in Asia and North America from Zizaniinae [5,6]. Besides rice (*Oryza sativa*), many other close relatives in this tribe could provide valuable genetic resources for rice breeding and improvement. For example, *Z. latifolia* had been used for generating introgression lines for potential breeding applications [7,8]. Furthermore, with the development of high throughput sequencing technologies, more species in the *Oryza* genus have been completed their whole genomes [9-13]. These progresses have made the tribe Oryzeae as an ideal system to study genomic evolution [14-16].

Now, in Oryzinae, numerous plastid genomes have been completely sequenced [17,18], however, the finished plastomes from subtribe of Zizaniinae are limited. *Zizania* is a small genus only comprised of four species including *Z. palustris*, *Z. aquatica*, and *Z. texana* native to North America, and one species *Z. latifolia* widely distributed in Eastern Asia [19]. As the only species native to Asia, *Z. latifolia*, named Manchurian wild rice, is used as a food plant with both the stem and grain being consumed in several Asian countries. As a relative of rice, *Zizania* could be an important genetic resource in rice breeding and genetic transformation [7]. Plastid transformation can accelerate genetic modifications, and therefore, the finished plastomes are particularly important genetic resource in developing improved rice varieties. Plastomes have been shown to be a very effective targets in genomic transformation with fewer concerns regarding gene flow into wild or non-transformed individuals [20].

Plastomes, similar to animal mitochondrial genomes, maintain a conserved circular double-stranded DNA structure, with sizes ranging

from 115 to 165 kb in land plants [21,22], and stable gene content and order [23]. These features of plastomes and the decreasing cost of high-throughput sequencing technologies have led to an increase in the number of completed plastomes. At present, more than 900 species' completed plastomes are currently available in the NCBI database (<http://www.ncbi.nlm.nih.gov/genomes/>). Besides being an effective tool for genetic transformation [24,25], complete plastomes are also important resources to explore the genetic and evolutionary variation between plant groups. For example, as a primary source for plant molecular systematic and taxonomic studies, plastomes have provided strong phylogenetic signals at multiple levels of inquiry [17,26-28]. For instance, the markers from plastomes have been used to explore the biogeographical relationships among plant populations [27,29] and for DNA barcoding [24,25].

In this study, we combined the traditional polymerase chain reaction (PCR) and Sanger sequencing methods to fully sequence and assemble the whole plastome of *Z. latifolia*. We conducted comprehensive comparisons with *Z. aquatica* [30] and other rice species to infer what evolutionary changes have occurred within the Oryzeae tribe. The finished plastome of *Z. latifolia* will provide important genetic information for plastid transformation involved with crop improvement.

***Corresponding author:** Zhiqiang Wu, Department of Biology, Colorado State University, Fort Collins, Colorado, USA, E-mail: wu.zhiqiang.1020@gmail.com

Cuihua Gu, School of Landscape and Architecture, Zhejiang Agriculture and Forestry University, Hangzhou 311300, P.R. China, E-mail: gu_cuihua@126.com

Received October 26, 2015; Accepted November 13, 2015; Published November 17, 2015

Citation: Wu Z, Gu C, Tembrock LR, Ge S (2015) Limited Polymorphisms between Two Whole Plastid Genomes in the Genus *Zizania* (Zizaniinae). J Proteomics Bioinform 8: 253-259. doi:10.4172/jpb.1000377

Copyright: © 2015 Wu Z, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Methods and Materials

Complete plastid genome sequence of *Zizania latifolia*

Fresh leaves of *Zizania latifolia* were collected from plants grown in the greenhouse of the Institute of Botany of the Chinese Academy of Sciences in Beijing. Total cellular DNA was extracted using the cetyltrimethyl ammonium bromide (CTAB) method [31] and purified with phenol extraction to remove proteins. With PCR amplification and Sanger sequencing methods, we completely sequenced the plastid genome (plastome) of *Z. latifolia*. For sequencing *Z. latifolia*, a full set of primers was designed based on two previously sequenced bamboo plastomes [17,32]. PCR amplification and purification of the products were performed, as described in Tang et al. [3]. The purified products were directly sequenced on an ABI 3730 automated sequencer (Applied Biosystems, Foster City, CA, USA). The sequences were assembled with the ContigExpress program from the Vector NTI Suite 6.0 (Informax Inc., North Bethesda, MD).

Plastid genome annotation and drawing

The fully assembled plastid genome of *Z. latifolia* was annotated using the DOGMA (Dual Organellar GenoMe Annotator, [33]). The first draft annotation from the DOGMA output was subsequently

manually inspected and adjusted for accurate assessment of the start and stop codons and the exon–intron boundaries of genes. Both tRNA and rRNA genes were identified by BLASTN searches against the database of chloroplast genomes from NCBI. The final annotations were plotted as a circular genome by the bioinformatics tools circos 0.67 [34] (Figure 1). For comparison, we download the published plastid genomes from five species in Oryzae and one bamboo species [17,35].

Dynamic variation of junction regions in plastid genome

Based on the configuration of the plastid genome with the two inverted repeat regions (IR_A and IR_B) and the two single copy regions (LSC and SSC) [36,37], four junctions named as J_{LA}, J_{LB}, J_{SA}, and J_{SB} are located between the two single copy (LSC and SSC) regions and the two IRs (IR_A and IR_B) [38]. The dynamic variation at the four junction regions (J_{LA}, J_{LB}, J_{SA}, and J_{SB}) can contribute to the size variation of plastid genomes [39,40]. The detailed IR border positions and the distance with adjacent genes among seven grass species plastomes (*Oryza sativa* ssp. *japonica* [AY522330], *O. australiensis* [KJ830774], *Leersia tisserantii* [JN415112], *Z. latifolia* [in this study], *Z. aquatica* [KJ870999], *Rhynchosyza subulata* [JN415114], *Phyllostachys propinqua* [JN415113]) were therefore compared in this work.

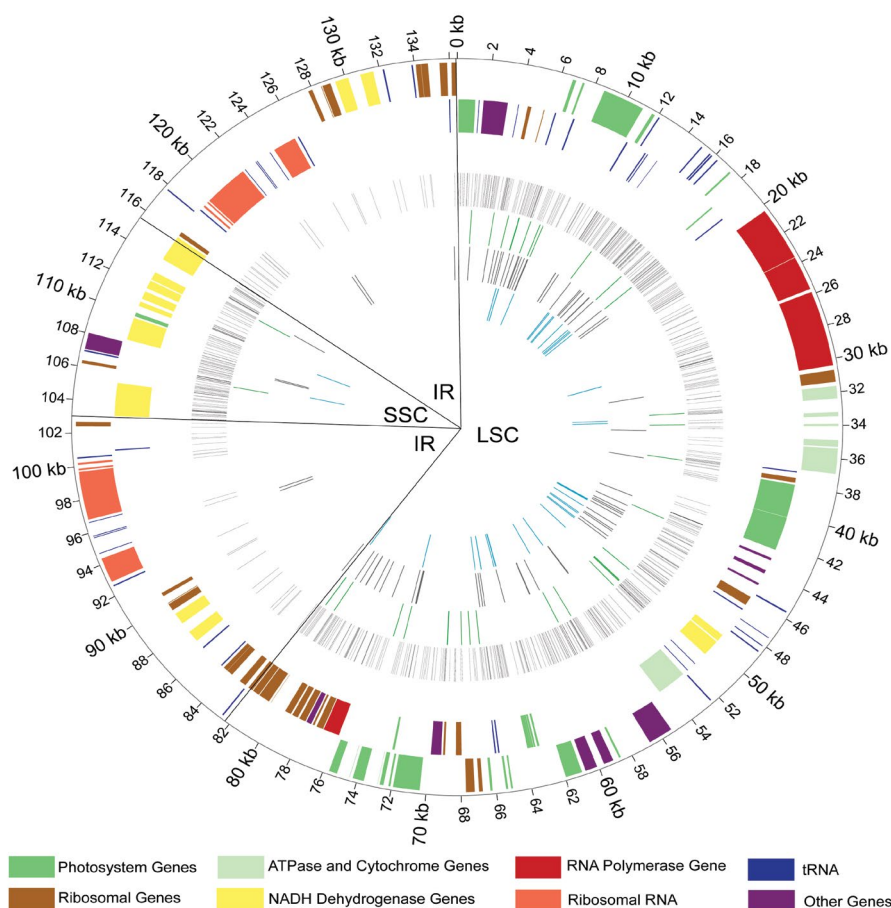


Figure 1: Chloroplast genome information and variation maps of *Z. latifolia* (KT161956). From outside to inside, tracks depict: 1) coding genes on forward strand; 2) coding genes on reverse strand; 3) the number and distribution of substitutions (simply named as SNP) (grey bar color); 4) the number and distribution of non-repeat insertion/deletions (Indels) (green bar color); 5) the number and distribution of poly structures (grey bar color); 6) the number and distribution of repeat Indels (green bar color). The coding genes are colored based on functional group with different color codes at the bottom. Maps were generated with software Circos v0.67.

Polymorphism analysis from coding and non-coding regions

To comprehensively compare the complete plastomes of two *Zizania* species, we employed the mVISTA program under the Shuffle-LAGAN mode [41] to detect whole genome variation with the plastome of *O. sativa* ssp. *japonica* (AY522330) as the reference. To identify the polymorphic regions that may be highly informative for phylogenetic or population genetic analyses from the two *Zizania* plastid genomes, we divided the whole plastid genome as coding and non-coding regions. Excluding the non-variation of the RNA genes, all protein coding genes were separately analyzed as coding regions, and all intron and intergenic regions were examined as non-coding regions (only the sequence longer than 200 bp were used in this analysis). These regions were extracted by using custom PERL scripts from the whole plastid genome and aligned using ClustalW under MEGA6 [42] and adjusted manually using the similarity criterion [43] for non-coding regions or preserving the reading frames for coding regions. The aligned sequences were used to calculate the sequence identity values (SI) by using BioEdit software [44] and the number of the polymorphic sites were reported from DnaSP v5.10 [45]. Only one part of the IRs regions was used in this process given the identical nature of the other repeat.

Phylogenetic analysis

To determine the phylogenetic relationships among the six Oryzaceae species and outgroup bamboo species, as whole plastomes alignment was used to build phylogenetic tree. Based on the collinearity of the whole plastid is extremely conserved among grass family [17,46], the whole plastomes of seven species were aligned in MAFFT v7.221 [47] under the FFT-NS-2 setting, followed by manual adjustment based on the similarity criterion [43]. Three different phylogenetic-inference methods were used to infer relationships: maximum parsimony analysis was implemented in PAUP* 4.0b10 [48], Bayesian inference (BI) in MrBayes 3.1.2 [49] and Neighbor-Joining (NJ) in MEGA6 [42] by applying the settings from Wu et al., [37].

Results and Discussion

Sequencing, assembly and annotation

Compared with the low-cost and high-throughput next generation sequencing (NGS) technologies, PCR and Sanger sequencing produced highly creditable results with fewer errors [37]. Meanwhile, some other features made plastid are easy to amplify including that the plastome is small but with abundant component in cellular DNA, and rates of nucleotide substitution are relatively slow [50]. By integrating the overlapped Sanger sequenced fragments (longer 100 bp) from PCR products, we successfully assembled the whole plastid genome for *Z. latifolia* with 136,461 bp in length by using the method from Wu et al., [32] (Figure 1). After the automatic annotation from DOGMA [33], comparative analyses, and manual verification, the annotated plastid genome of *Z. latifolia* was uploaded to GenBank with accession KT161956 (Figure 1). It is composed of two single-copy regions separated by a pair of inverted repeats (IR) of 20,878 bp each, which account for 30.60% of the whole plastid genome. The large single copy (LSC) and the small single copy (SSC) regions span 82,115 bp and 12,590 bp, respectively. The proportion of LSC and SSC length in the total plastid genome is 60.17% and 9.23% respectively, and those features were all extremely similar with the other six published plastomes (Table 1). The same gene content, gene number and gene order among the seven species reflects the highly conserved plastomes of Poaceae [46,51]. The size differences among plastomes are mainly a result of the variation within intergenic sequences.

Polymorphisms between the two *Zizania* plastomes

Discovery of polymorphisms between plastomes are the basis of marker selection in plant systematic or barcoding research [24,52], in addition, discovery of sites with limited variation could be the candidates of plastid transformation [20]. We conducted the following genome wide and local specific analysis to fully investigate the polymorphisms variation between two *Zizania* plastomes.

First, based on the conserved features of the land plant plastomes [22], the polymorphisms were discovered from the whole genome alignment of two *Zizania* plastomes and categorized as either

Subfamily	Species	Total size		LSC region		IR region		SSC region		Gene Content	GenBank Accession
		Length (bp)	GC (%)	Length (bp)	GC (%)	Length (bp)	GC (%)	Length (bp)	GC (%)		
Ehrhartoideae	<i>Oryza sativa</i> ssp. <i>japonica</i>	134,551	39.00	80,604	37.11	20,802	44.35	12,343	33.37	the same	AY522330
	<i>Oryza australiensis</i>	134,549	38.98	80,614	37.07	20,796	44.36	12,343	33.25	the same	GU592209
	<i>Leersia tisserantii</i>	136,550	38.88	81,865	37.01	21,329	44.05	12,027	33.23	the same	JN415112
	<i>Zizania latifolia</i>	136,461	39.00	82,115	37.13	20,878	44.42	12,590	33.18	the same	KT161956 ^a
	<i>Zizania aquatica</i>	136,364	39.02	82,013	37.14	20,879	44.41	12,593	33.31	the same	KJ870999
	<i>Rhynchosyza subulata</i>	136,303	39.00	82,029	37.14	20,840	44.36	12,594	33.40	the same	JN415114
Bambusoideae	<i>Phyllostachys propinqua</i>	139,704	38.88	83,227	36.96	21,800	44.23	12,877	33.14	the same	JN415113

^a Sequenced in this study.

Table 1: Comparison of major features of seven Poaceae plastome.

Type	Region	Coding Regions			Non-Coding Regions			Sum
SNP	LSC	189			379			568
	IR	8			38			46
	SSC	70			60			130
	Total	267			477			744
Indel	Region	Indel	Poly	Repeat	Indel	Poly	Repeat	
	LSC	0	0	1	28	63	32	124
	IR	0	0	0	0	6	0	6
	SSC	0	0	0	2	3	2	7
Total	0	0	1	30	72	34	137	

Table 2: The number of polymorphisms in different regions by comparison from two *Zizania* plastome.

substitutions (for simply named as SNPs) or insertions/deletions (Indels, Table 2). From the results (Table S1 and S2), we found that the large single copy (LSC) regions possessed the most variable sites (568 SNPs and 124 Indels), and the two inverted repeats (IR) contained the least variable sites (46 SNPs and 6 Indels). However, when we consider the number of variable sites per kilo base pairs (kbp), the small single copy (SSC) region was the highest with 10.3 substitutions/kbp, and 6.9 and 2.2 for the LSC and IR. This result was the same as

the evolutionary rate variation among the three regions [53]. When comparing polymorphisms between coding and non-coding regions, the non-coding regions contained the most variable sites in the LSC and IR regions, but in SSC, the number of SNPs was higher in coding regions. This result shows that in LSC and SSC regions, the rate of variation between coding and non-coding is different.

Second, for comparison the whole genome variation within the two *Zizania* plastomes, we also employed the mVISTA program

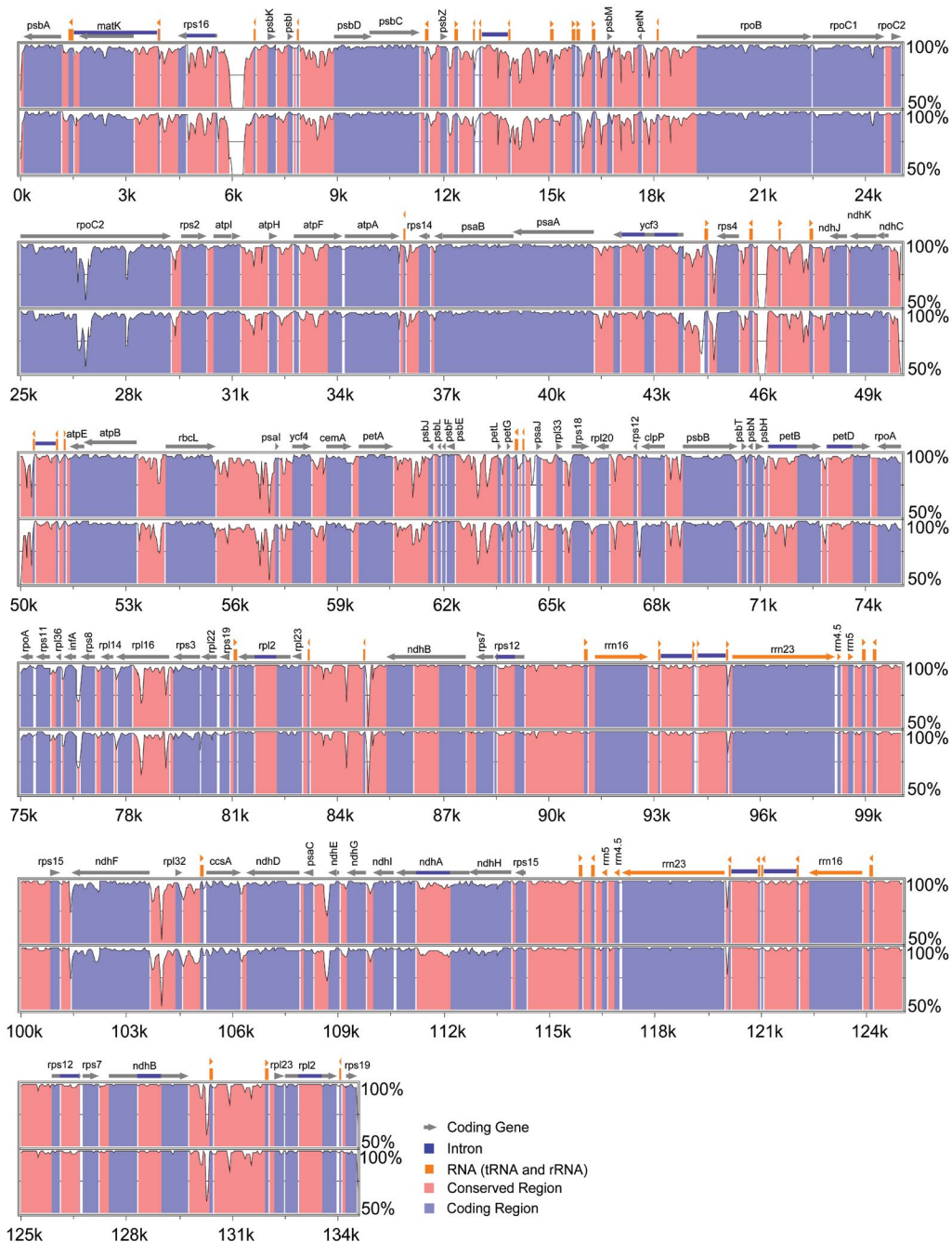


Figure 2: Identity plot that compares the chloroplast genomes of the two *Zizania* data sets used in this study with *O. sativa* ssp. *Japonica* (AY522330) as the reference sequence. The vertical scale indicates the percentage of identity to the reference, ranging from 50% to 100%. The horizontal axis indicates the coordinated base position within the chloroplast genome. Genome regions are color coded as protein-coding, RNA (rRNA and tRNA), intron, and conserved noncoding regions.

[41]. We used *O. sativa* ssp. *japonica* (AY522330) as the reference for these comparisons (Figure 2). In regards to structural differences, the organization of the plastome was conserved between the two *Zizania* species with no differing translocations or inversions found. The two IR regions were more conserved than the LSC and SSC regions with 99.8% similarity found between them, and the coding regions were more conserved than non-coding regions (by the view from different colors). The most variation from coding regions was found in the *rpoC2* gene (Figure 3). After the alignment of the amino acid (AA) sequence of *rpoC2*, the variable regions were focused around 650 to 820 AA. The length of this gene was different in all seven species surveyed and was mainly caused by size variation of repeat sequence within *rpoC2*. For the two *Zizania* species, there was one only 7 AA deletion around position 680 found in *Z. aquatica*. Among the non-coding regions (the pink portions in Figure 2), *ndhC-trnV*, *rps16-trnQ* and *ycf3-trnS* demonstrated the most divergent patterns between the two *Zizania* species.

Third, to examine the sequence identity (SI value) of the two *Zizania* plastomes, the whole plastome was divided into individual coding (gene regions) and non-coding (intergenic and intronic regions). Among 76 annotated protein coding genes (Table S3), which were the same in gene content as other grass species [17], the SI values for the two *Zizania* species were all larger than 98% with the average value for SI being 99.5% with 23 genes showing no variation. These results showed that all coding genes were extremely conserved between two *Zizania* species. However, the genes *matK*, *ndhF* and *rpoC2* possessed the lowest SI values, and these regions have been used as universal markers in barcoding and phylogenetic studies [54], which also show phylogenetic signal at lower taxonomic levels. For the SI values from 126 non-coding regions (only one copy of IR regions was used) (Table S4), we found that the average SI values across all non-coding regions

longer than 200 bp was 97.9% and was only a little lower than coding regions. Only three regions (*ndhC-trnV*, *rps16-trnQ* and *ycf3-trnS*) had SI values lower than 90%. *Rps16-trnQ* had the greatest number of polymorphic sites and as such was also an important marker for plant systematic studies [54].

Conserved variation of two *Zizania* plastome junction boundaries

Based on the quadripartite structure of plastomes from most land plants [22] (Figure 4A), four junctions (J_{LA} , J_{LB} , J_{SA} , and J_{SB}) were named between the two IRs (IRa and IRb) and the two single copy (LSC and SSC) regions (Figure 4C) [38]. It has already been shown that the dynamic variation of IR regions can contribute to the size variation between plastomes [22] and as such provide useful phylogenetic signals [38]. By combining the phylogenetic relationships among the seven grass species plastomes species (Figure 4B) and applying the distance between border position and the adjacent genes, we were able to compare the gene distance in a phylogenetic context (Figure 4C). By comparison with the variation from the *Oryza* genus [37], the two species of *Zizania* retained the same distance around the four junctions (Figure 3B). This indicates that the distance between gene and junction border is more conserved in *Zizania* than *Oryza*. We also examined the difference at gene distance to junction between the two subtribes. The distance of J_{SA} and J_{SB} reflects the variation of the two *Oryzeae* subtribes from this plastome feature. In subtribe *Zizaniinae*, the two genes distance from J_{SB} is 319 bp (*rps15*) and 89-93 bp (*ndhF*). However, in subtribe *Oryzinae*, those distances were 301 bp and 41-42 bp (except for *L. tisserantii*). This feature could be used as a marker to separate the two subtribes. In total, the junction in J_{LA} and J_{LB} are more conserved than J_{SA} and J_{SB} .

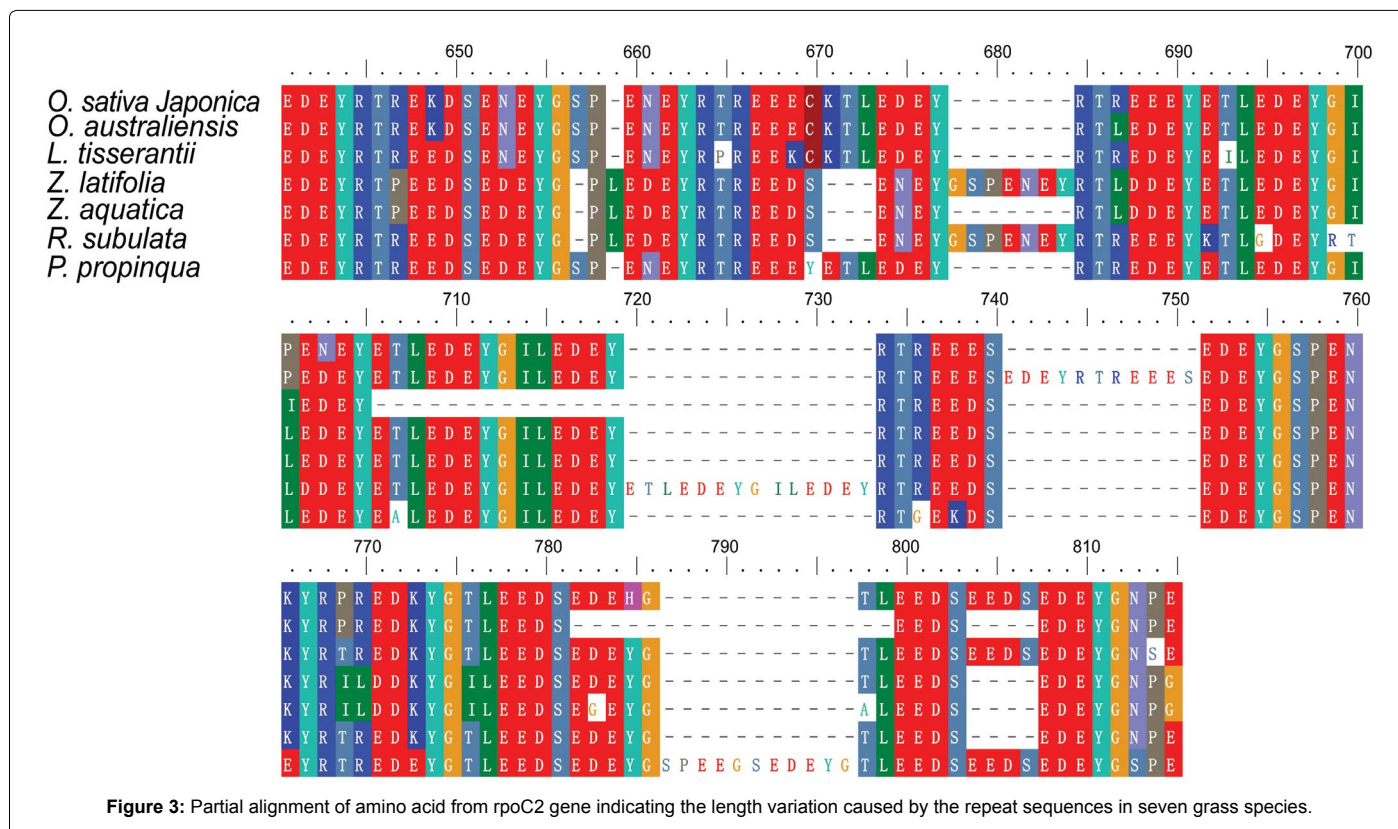
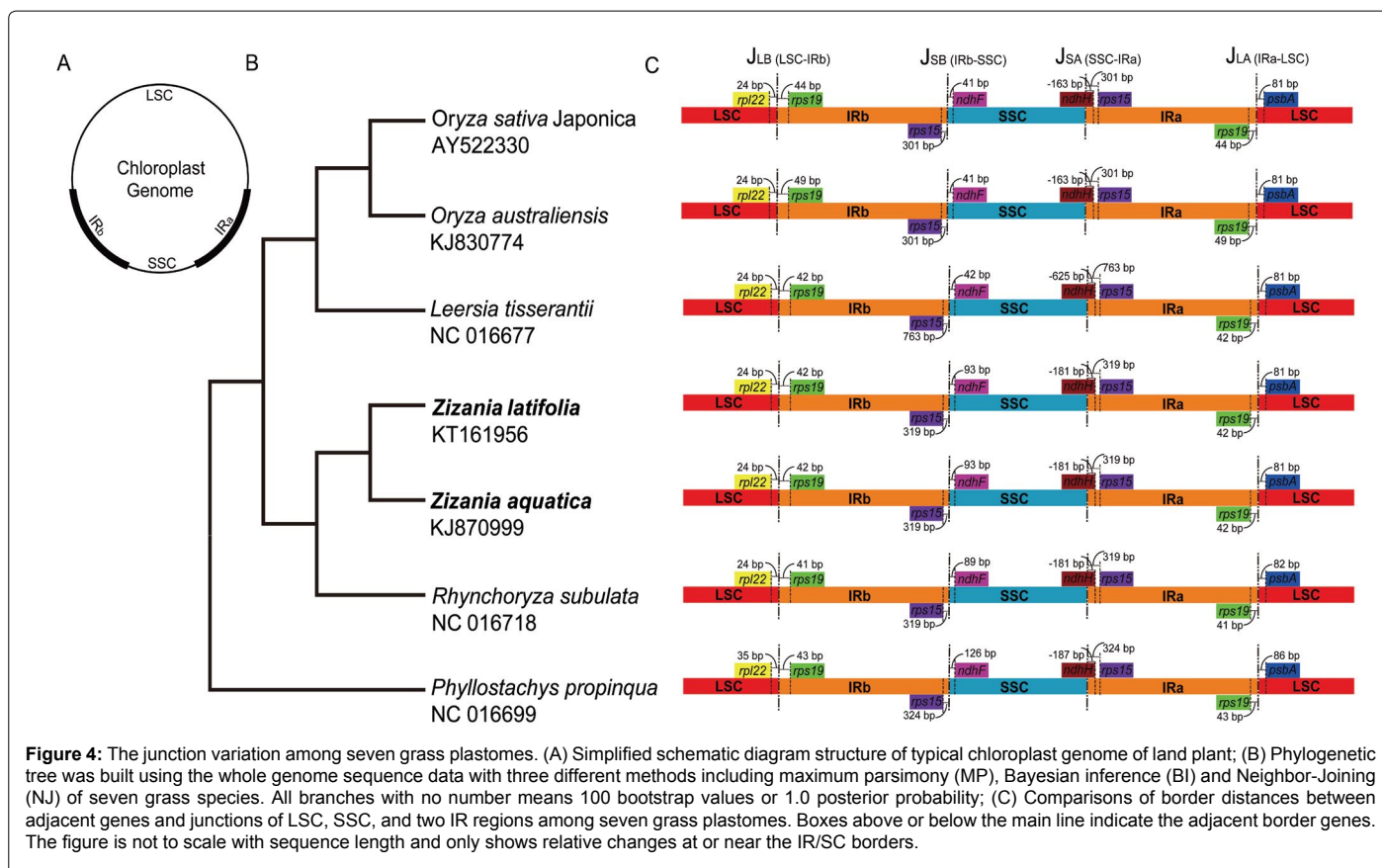


Figure 3: Partial alignment of amino acid from *rpoC2* gene indicating the length variation caused by the repeat sequences in seven grass species.



Acknowledgement

This work was supported by the National Natural Science Foundation of China (30990240). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Reference

- Ge S, Li A, Lu BR, Zhang SZ, Hong DY (2002) A phylogeny of the rice tribe Oryzaceae (Poaceae) based on matK sequence data. *Am J Bot* 89: 1967-1972.
- Guo YL, Ge S (2005) Molecular phylogeny of Oryzaceae (Poaceae) based on DNA sequences from chloroplast, mitochondrial, and nuclear genomes. *Am J Bot* 92: 1548-1558.
- Tang L, Zou XH, Achoundong G, Potgieter C, Second G, et al. (2010) Phylogeny and biogeography of the rice tribe (Oryzaceae): evidence from combined analysis of 20 chloroplast fragments. *Mol Phylogenet Evol* 54: 266-277.
- Tang L, Zou XH, Zhang LB, Ge S4 (2015) Multilocus species tree analyses resolve the ancient radiation of the subtribe Zizaniinae (Poaceae). *Mol Phylogenet Evol* 84: 232-239.
- Johnson E (1969) Archeological evidence for utilization of wild rice. *Science* 163: 276-277.
- Vaughan DA (1994) The wild relatives of rice: a genetic resources handbook. 1994.
- Dong ZY, Wang YM, Zhang ZJ, Shen Y, Lin XY, et al. (2006) Extent and pattern of DNA methylation alteration in rice lines derived from introgressive hybridization of rice and *Zizania latifolia* Griseb. *Theor Appl Genet* 113: 196-205.
- Dong Z, Wang H, Dong Y, Wang Y, Liu W, et al. (2013) Extensive microsatellite variation in rice induced by introgression from wild rice (*Zizania latifolia* Griseb.). *PLoS One* 8: e62317.
- Goff SA, Ricke D, Lan TH, Presting G, Wang R, et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296: 92-100.
- Yu J, Hu S, Wang J, Wong GK, Li S, et al. (2002) A draft sequence of the rice

genome (*Oryza sativa* L. ssp. *indica*). *Science* 296: 79-92.

- Kim H, Hurwitz B, Yu Y, Collura K, Gill N, et al. (2008) Construction, alignment and analysis of twelve framework physical maps that represent the ten genome types of the genus *Oryza*. *Genome Biol* 9: R45.
- Wang M, Yu Y, Haberer G, Marri PR, Fan C, et al. (2014) The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication. *Nat Genet* 46: 982-988.
- Zhang QJ, Zhu T, Xia EH, Shi C, Liu YL, et al. (2014) Rapid diversification of five *Oryza* AA genomes associated with rice adaptation. *Proc Natl Acad Sci U S A* 111: E4954-4962.
- Zou XH, Zhang FM, Zhang JG, Zang LL, Tang L, et al. (2008) Analysis of 142 genes resolves the rapid diversification of the rice genus. *Genome Biol* 9: R49.
- Ai B, Wang ZS, Ge S (2012) Genome size is not correlated with effective population size in the *Oryza* species. *Evolution* 66: 3302-3310.
- Jacquemin J, Ammiraju JS, Haberer G, Billheimer DD, Yu Y, et al. (2014) Fifteen million years of evolution in the *Oryza* genus shows extensive gene family expansion. *Mol Plant* 7: 642-656.
- Wu ZQ, Ge S (2012) The phylogeny of the BEP clade in grasses revisited: evidence from the whole-genome sequences of chloroplasts. *Mol Phylogenet Evol* 62: 573-578.
- Wambugu PW, Brozynska M, Furtado A, Waters DL, Henry RJ1 (2015) Relationships of wild and domesticated rices (*Oryza* AA genome species) based upon whole chloroplast genome sequences. *Sci Rep* 5: 13957.
- Xu XW, Wu JW, Qi MX, Lu QX, Lee PF, et al. (2015) Comparative phylogeography of the wild-rice genus *Zizania* (Poaceae) in eastern Asia and North America. *Am J Bot* 102: 239-247.
- Bock R1 (2015) Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. *Annu Rev Plant Biol* 66: 211-241.
- Palmer JD (1985) Comparative organization of chloroplast genomes. *Annu Rev Genet* 19: 325-354.

22. Ravi V, Khurana JP, Tyagi AK, Khurana P (2008) An update on chloroplast genomes. *Plant Syst Evol* 271: 101-122.
23. Wicke S, Schneeweiss GM, dePamphilis CW, Muller KF, Quandt D (2011) The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol Biol* 76: 273-297.
24. Group CPBOL, Li DZ, Gao LM, Li HT, Wang H, et al. (2011) Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proc Natl Acad Sci U S A* 108: 19641-19646.
25. Day A, Goldschmidt-Clermont M (2011) The chloroplast transformation toolbox: selectable markers and marker removal. *Plant Biotechnol J* 9: 540-553.
26. Wang L, Qi X, Xiang Q, Heinrichs J, Schneider H, Zhang X (2010) Phylogeny of the paleotropical fern genus *Lepisorus* (*Polypodiaceae*, *Polypodiopsida*) inferred from four chloroplast DNA regions. *Mol Phylogenet Evol* 54: 211-225.
27. Wang L, Wu ZQ, Bystrakova N, Ansell SW, Xiang QP, et al. (2011) Phylogeography of the Sino-Himalayan fern *Lepisorus clathratus* on "the roof of the world". *PLoS One* 6: e25896.
28. Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, et al. (2007) Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc Natl Acad Sci U S A* 104: 19369-19374.
29. Wang L, Schneider H, Zhang XC, Xiang QP (2012) The rise of the Himalaya enforced the diversification of SE Asian ferns by altering the monsoon regimes. *BMC Plant Biol* 12: 210.
30. Wysocki WP, Clark LG, Attigala L, Ruiz-Sanchez E, Duvall MR5 (2015) Evolution of the bamboos (*Bambusoideae*; *Poaceae*): a full plastome phylogenomic analysis. *BMC Evol Biol* 15: 50.
31. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19: 11-15.
32. Wu FH, Kan DP, Lee SB, Daniell H, Lee YW, et al. (2009) Complete nucleotide sequence of *Dendrocalamus latiflorus* and *Bambusa oldhamii* chloroplast genomes. *Tree Physiol* 29: 847-856.
33. Wyman SK, Jansen RK, Boore JL (2004) Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* 20: 3252-3255.
34. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, et al. (2009) Circos: an information aesthetic for comparative genomics. *Genome Res* 19: 1639-1645.
35. Wu ZQ, Ge S (2014) The whole chloroplast genome of wild rice (*Oryza australiensis*). *Mitochondrial DNA* 1-2.
36. Palmer JD (1983) Chloroplast DNA exists in two orientations. *Nature* 301: 92-93.
37. Wu Z, Tembrock LR, Ge S3 (2015) Are differences in genomic data sets due to true biological variants or errors in genome assembly: an example from two chloroplast genomes. *PLoS One* 10: e0118019.
38. Wang RJ, Cheng CL, Chang CC, Wu CL, Su TM, et al. (2008) Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evol Biol* 8: 36.
39. Palmer JD, Nugent JM, Herbon LA (1987) Unusual structure of geranium chloroplast DNA: A triple-sized inverted repeat, extensive gene duplications, multiple inversions, and two repeat families. *Proc Natl Acad Sci U S A* 84: 769-773.
40. Nazareno AG, Carlsen M, Lohmann LG1 (2015) Complete Chloroplast Genome of *Tanaecium tetragonolobum*: The First Bignoniaceae Plastome. *PLoS One* 10: e0129930.
41. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I (2004) VISTA: computational tools for comparative genomics. *Nucleic Acids Res* 32: W273-279.
42. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30: 2725-2729.
43. Simmons MP1 (2004) Independence of alignment and tree search. *Mol Phylogenet Evol* 31: 874-879.
44. Eckert AJ, Hall BD (2006) Phylogeny, historical biogeography, and patterns of diversification for *Pinus* (Pinaceae): phylogenetic tests of fossil-based hypotheses. *Mol Phylogenet Evol* 40: 166-182.
45. Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.
46. Cotton JL, Wysocki WP, Clark LG, Kelchner SA, Pires JC, et al. (2015) Resolving deep relationships of PACMAD grasses: a phylogenomic approach. *BMC Plant Biol* 15: 178.
47. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30: 772-780.
48. Swofford DL (2002) PAUP*: phylogenetic analysis using parsimony (and other methods).
49. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61: 539-542.
50. Clegg MT, Gaut BS, Learn GH Jr, Morton BR (1994) Rates and patterns of chloroplast DNA evolution. *Proc Natl Acad Sci U S A* 91: 6795-6801.
51. Michelangeli FA, Davis JI, Stevenson DW (2003) Phylogenetic relationships among Poaceae and related families as inferred from morphology, inversions in the plastid genome, and sequence data from the mitochondrial and plastid genomes. *Am J Bot* 90: 93-106.
52. Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, et al. (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am J Bot* 92: 142-166.
53. Walker JF, Zanis MJ, Emery NC (2014) Comparative analysis of complete chloroplast genome sequence and inversion variation in *Lasthenia burkei* (Madieae, Asteraceae). *Am J Bot* 101: 722-729.
54. Dong W, Liu J, Yu J, Wang L, Zhou S (2012) Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS One* 7: e35071.