University of Pennsylvania
ScholarlyCommons

Departmental Papers (Dental) Penn Dental Medicine

6-2009

Complete Nucleotide Sequence of Dendrocalamus Latiflorus and Bambusa Oldhamii Chloroplast Genomes

F. H. Wu

D. P. Kan

S. B. Lee

Henry Daniell
University of Pennsylvania

Y. W. Lee

See next page for additional authors

Follow this and additional works at: https://repository.upenn.edu/dental_papers

Part of the Dentistry Commons

Recommended Citation

Wu, F. H., Kan, D. P., Lee, S. B., Daniell, H., Lee, Y. W., Lin, C. C., Lin, N. S., & S., C. (2009). Complete Nucleotide Sequence of Dendrocalamus Latiflorus and Bambusa Oldhamii Chloroplast Genomes. *Tree Physiology, 29* (6), 847-856. http://dx.doi.org/10.1093/treephys/tpp015

At the time of publication, author Henry Daniell was affiliated with the University of Central Florida. Currently, he is a faculty member at the School of Dental Medicine at the University of Pennsylvania.

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/dental_papers/217
For more information, please contact repository@pobox.upenn.edu.

Complete Nucleotide Sequence of Dendrocalamus Latiflorus and Bambusa Oldhamii Chloroplast Genomes

Abstract

Although bamboo is one of the most important woody crops in Asia, information on its genome is still very limited. To investigate the relationship among Poaceae members and to understand the mechanism of albino mutant generation in vitro, the complete chloroplast genome of two economically important bamboo species, *Dendrocalamus latiflorus* Munro and *Bambusa oldhamii* Munro, was determined employing a strategy that involved polymerase chain reaction (PCR) amplification using 443 novel primers designed to amplify the chloroplast genome of these two species. The lengths of the *B. oldhamii* and *D. latiflorus* chloroplast genomes are 139,350 and 139,365 bp, respectively. The organization structure and the gene order of these two bamboos are identical to other members of Poaceae. Highly conserved chloroplast genomes of Poaceae facilitated sequencing by the PCR method. Phylogenetic analysis using both chloroplast genomes confirmed the results obtained from studies on chromosome number and reproductive organ morphology. There are 23 gaps, insertions/deletions > 100 bp, in the chloroplast genomes of 10 genera of Poaceae compared in this study. The phylogenetic distribution of these gaps corresponds to their taxonomic placement. The sequences of these two chloroplast genomes provide useful information for studying bamboo evolution, ecology and biotechnology.

Keywords

Bambusoideae, biotechnology, phylogenetic analysis, Poaceae

Disciplines

Dentistry

Comments

At the time of publication, author Henry Daniell was affiliated with the University of Central Florida. Currently, he is a faculty member at the School of Dental Medicine at the University of Pennsylvania.

Author(s)

F. H. Wu, D. P. Kan, S. B. Lee, Henry Daniell, Y. W. Lee, C. C. Lin, N. S. Lin, and C. S.

This journal article is available at ScholarlyCommons: https://repository.upenn.edu/dental_papers/217

NIH Public Access
Author Manuscript
Tree Physiol. Author manuscript; available in PMC 2009 December 1.

Published in final edited form as:
Tree Physiol. 2009 June ; 29(6): 847–856. doi:10.1093/treephys/tpp015.

Complete nucleotide sequence of *Dendrocalamus latiflorus* and *Bambusa oldhamii* chloroplast genomes

F.-H. WU[1,5], **D.-P. KAN**[1,2,5], **S.-B. LEE**[3], **H. DANIELL**[3], **Y.-W. LEE**[1], **C.-C. LIN**[2], **N.-S. LIN**[1,4], and **C.-S. LIN**[1,6]

[1] Agricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan

[2] Department of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, Taipei, Taiwan

[3] Department of Molecular Biology and Microbiology, College of Medicine, University of Central Florida, Orlando, FL 32816-2364, USA

[4] Institute of Plant and Microbial Biology, Academia Sinica, Taipei, Taiwan

Summary

Although bamboo is one of the most important woody crops in Asia, information on its genome is still very limited. To investigate the relationship among Poaceae members and to understand the mechanism of albino mutant generation in vitro, the complete chloroplast genome of two economically important bamboo species, *Dendrocalamus latiflorus* Munro and *Bambusa oldhamii* Munro, was determined employing a strategy that involved polymerase chain reaction (PCR) amplification using 443 novel primers designed to amplify the chloroplast genome of these two species. The lengths of the *B. oldhamii* and *D. latiflorus* chloroplast genomes are 139,350 and 139,365 bp, respectively. The organization structure and the gene order of these two bamboos are identical to other members of Poaceae. Highly conserved chloroplast genomes of Poaceae facilitated sequencing by the PCR method. Phylogenetic analysis using both chloroplast genomes confirmed the results obtained from studies on chromosome number and reproductive organ morphology. There are 23 gaps, insertions/deletions > 100 bp, in the chloroplast genomes of 10 genera of Poaceae compared in this study. The phylogenetic distribution of these gaps corresponds to their taxonomic placement. The sequences of these two chloroplast genomes provide useful information for studying bamboo evolution, ecology and biotechnology.

Keywords

Bambusoideae; biotechnology; phylogenetic analysis; Poaceae

---

[6]Corresponding author (cslin99@gate.sinica.edu.tw).
[5]These two authors contributed equally

Authors' contributions

D.P.K. and F.H.W. participated in the design of the study, designed the primers for amplification and sequencing, and performed the PCR. H.D. and S.B.L. participated in the annotation, drew chloroplast genome maps and prepared the manuscript. Y.W.L. participated in phylogenetic analysis. C.C.L. and N.S.L. participated in the manuscript preparation and experiment design. C.S.L. participated in the design of the study, contig assembly, complete chloroplast genome comparison and gap identification, developed the figures and prepared the manuscript. All authors have read and approved the final manuscript.

Supplementary Data
Supplementary data for this article are available at *Tree Physiology* Online.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Introduction

The group of woody, perennial, evergreen plants, commonly referred to as bamboo, belong to the grass family (Poaceae, Bambusoideae, Bambuseae). Bamboos have considerable economic and cultural significance in most areas of East Asia and Southeastern Asia. There are two types of bamboo: pachymorph and leptomorph rhizome (Stapleton 1997). In Taiwan and Southeastern Asia, the pachymorph rhizome is the most abundant type. In terms of economic and environmental impacts, *Bambusa* and *Dendrocalamus* are the two most important pachymorph rhizome genera. *Bambusa oldhamii* Munro, known as 'green timber bamboo', is a well-known semitropical clumping bamboo species native of Taiwan and southern China (Hsu et al. 2000). Very young green bamboo shoots are edible and considered to be a high-price culinary delicacy in Taiwan. *Dendrocalamus latiflorus* Munro is an evergreen species locally known as 'Taiwan giant bamboo'. This giant bamboo, which forms abundant forests in Taiwan, is a valuable natural resource, with its mature culms used as building material and its young shoots cultivated extensively for human consumption (Lin et al. 2007). Flowering is very difficult to induce in bamboo, thereby limiting its commercial cultivation. In addition, the quality of bamboo stands is reduced because of infection by the Bamboo Mosaic Virus. This has led to the establishment of reliable tissue culture systems for bamboo (Lin and Chang 1998, Hsu et al. 2000, Lin et al. 2007).

Photosynthesis occurs in specialized organelles called chloroplasts that are found in all higher plant cells. Because chloroplasts have their own genome and protein-synthesizing machinery, they are inherited independent of the nuclear genome (Martin et al. 1998). Consequently, most cases of spontaneous generation of albino mutant lines during plant cell culture are considered to be the result of mutation(s) in the chloroplast genome. One such example is the chloroplast genome-deleted albino mutant isolated during wheat anther culture (Day and Ellis 1984). Spontaneous deletions in the chloroplast genome can also occur in vivo. In the tobacco chloroplast genome, there are two repeats on either side of the large subunit of the ribulose-bisphosphate carboxylase (*rbcL*) gene. Kode et al. (2006) obtained cells with an albino pattern due to recombination, followed by deletion of the *rbcL* gene in some of the mesophyll cells. Bamboo tissue culture systems have also spontaneously produced albino plants with deletion in chloroplast genomes (Liu et al. 2007, Lin et al. 2008). The deletion in the chloroplast genome of these albino mutants has been confirmed by polymerase chain reaction (PCR) analysis; however, the mechanisms of albino mutant generation are still unknown, although the structure of the genome and repeat sequences facilitate spontaneous generation of albino mutants in other systems.

Because the chloroplast has its own protein synthesizing machinery and its genome is highly conserved, it is an ideal experimental tool for investigating evolutionary relationships (Palmer 1985). An important study by the Grass Phylogeny Working Group (2001) determined the relationships among grasses using a large and diverse assembly of species and by comparing three chloroplasts and three nuclear markers. Based on the results, the existence of two major lineages were highlighted, the BEP (Bambusoideae, Ehrhartoideae, Pooideae) clade and the PACCAD (Panicoideae, Aristidoideae, Centothecoideae, Chloridoideae, Arundinoideae, Danthonioideae) clade, which together comprise the majority of grasses. Many of the chloroplast genome sequences of two subfamilies of the BEP clade, Ehrhartoideae (rice) and Pooideae (wheat, barley, etc.), have been published (Ogihara et al. 2002, Tang et al. 2004, Saski et al. 2007, Bortiri et al. 2008). However, there is no information available on the subfamily Bambusoideae. It is evident that many additional chloroplast genome sequences are needed to provide sufficient taxon sampling to enable the generation of a family-wide phylogeny of grasses based on whole genomes (Saski et al. 2007). As bamboo is one of the most important Poaceae species in the wild, more information on the chloroplast genome will be useful for further investigations.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Chloroplast genetic engineering offers a number of unique advantages including high levels of transgene expression in healthy plants (up to 46% of leaf protein, DeCosa et al. 2001), multi-gene engineering in a single transformation event, transgene containment via maternal inheritance, and lack of gene silencing and position effect due to site-specific transgene integration (Quesada-Vargas et al. 2005, Daniell 2007, Verma and Daniell 2007). Because of these advantages, several transgenes have been integrated into chloroplast genomes to confer valuable agronomic traits, including herbicide/insect resistance, disease resistance, drought tolerance, salt tolerance and phytoremediation (McBride et al. 1995, Daniell et al. 1998, DeGray et al. 2001, Lee et al. 2003, Ruiz et al. 2003, Kumar et al. 2004). Moreover, chloroplast genetic engineering has been used to study chloroplast biogenesis and function, revealing the mechanisms of DNA replication origins, intron maturases, translation elements and proteolysis, import of proteins and several other biological processes (Grevich and Daniell 2005, Verma and Daniell 2007). Chloroplast has been used as a bioreactor to produce vaccine antigens and biopharmaceuticals (Staub et al. 2000, Leelavathi and Reddy 2003, Koya et al. 2005, Tregoning et al. 2005, Arlen et al. 2007, 2008, Verma and Daniell 2007). Although many successful examples of plastid engineering have laid a solid foundation for various biotechnology applications, this approach has not been extended to grasses; however, some progress has been made in rice (Lee et al. 2006*a*). The lack of complete chloroplast genome sequences is still one of the major limitations for extending this technology to other crops. Until 2004, only a handful of crop chloroplast genomes were sequenced. However, with the advancement of new sequencing technologies, in the past few years, > 30 crop chloroplast genomes have been sequenced (for a current list of complete annotated chloroplast genome sequences and their accession numbers see Verma et al. 2008).

There are many different strategies and tools used for sequencing the chloroplast genome. When bacterial artificial chromosome (BAC) libraries are available, clones containing chloroplast genomes could be identified using chloroplast genes as probes. Several crop chloroplast genomes have been sequenced using BAC libraries (Saski et al. 2005, Daniell et al. 2006, Jansen et al. 2006, Daniell et al. 2008). When BAC libraries are not available, intact chloroplasts have been isolated and whole chloroplast genome has been amplified using the rolling circle amplification method (Jansen et al. 2005, Lee et al. 2006*b*, Ruhlman et al. 2006, Samson et al. 2007). A practical, easy-to-use, reliable method for isolating and sequencing the chloroplast genome of the target species is an important goal. The slow evolutionary rate and highly conserved nature of the chloroplast genome and gene order has enabled researchers to design universal primers for the purpose of working across species' boundaries (Demesure et al. 1995, Dumolin-Lapegue et al. 1997). Grivet et al. (2001) published 38 primer pairs for amplifying the large single copy (LSC) region in angiosperms, but these primers are not genome wide. In 2005, a method called amplification, sequencing and annotation of plastomes (ASAP) was developed. Dhingra and Folta (2005) reported that 27 primer pairs derived from the inverted repeats (IRs) of five sequenced eudicots were sufficient to amplify the corresponding ~ 30-kb region in commercial strawberry. This result indicated that the chloroplast genome can be sequenced by PCR approach. Logacheva et al. (2008) also used the ASAP method to determine the chloroplast genome of *Fagopyrum esculentum* Moench. ssp. *ancestrale*. In an attempt to centralize all available information, Heinze (2007) developed a database on primers for chloroplast genomes and designed additional primers to fill gaps where the primer information is still limited. At present, there are 733 primers in this database. In a number of recent studies, Lin et al. used the PCR method to identify the chloroplast genome deletion in an albino bamboo mutant (Liu et al. 2007, Lin et al. 2008), thereby demonstrating a PCR-based method for sequencing the bamboo chloroplast genome. We report here, for the first time, the sequencing of the chloroplast genomes of two economically important bamboo species, *D. latiflorus* and *B. oldhamii*, using an easy PCR-based method.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Materials and methods

Plant materials

The leaves of *B. oldhamii* and *D. latiflorus* used in this study were obtained from the greenhouse in Academia Sinica, Taipei, Taiwan. These bamboos originated from a local variety collected from Tainan, Taiwan.

DNA purification, primer design and genomic PCR

For the chloroplast genomic PCR analysis, total genomic DNA from greenhouse-grown plants was isolated using a urea extraction buffer system (Sheu et al. 1996). Figure 1 shows the strategy used for chloroplast genome sequencing. The coding regions of rice chloroplast genome are used as the templates for primer design. The primer sequences were submitted to the *Saccharomyces* genome database (http://www.yeastgenome.org/) and a melting temperature of 55 °C was used to select suitable primers. Each DNA fragment was then amplified using specific primers (Online Supplementary Data). The PCR amplification program consisted of one cycle at 94 °C for 5 min; 25 cycles at 94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min and a final cycle at 72 °C for 5 min; the PCR products were sequenced. DNA sequencing was carried out with the Big-Dye Terminator Cycle Sequencing kit using an ABI Prism 3700 DNA analyzer (Applied Biosystems, Foster City, CA). The sequences were compared with the chloroplast genome of rice using the VectorNTI, AlignX software program (Version 7.0; Invitrogen, Carlsbad, CA; parameters: overlap: 30; identity: 0.95 and cutoff score: 40). Those sequences that did not match the rice chloroplast genome were deleted, and the gaps were filled by the new primers that were designed on the basis of sequences obtained from the PCR products.

Contig assembly and annotation

VectorNTI contig express was used to assemble contigs (parameters: overlap: 30; identity: 0.95 and cutoff score: 40). Annotation of both chloroplast genomes was performed using DOGMA (Dual Organellar GenoMe Annotator, Wyman et al. 2004, http://dogma.cc.bb.utexas.edu). This program uses a FASTA-formatted input file of the complete genomic sequences and identifies putative protein-coding genes by performing BLASTX searches against a custom database of previously published chloroplast genomes. We selected putative start and stop codons for each protein-coding gene and intron and exon boundaries for intron-containing genes. Both tRNAs and rRNAs are identified by BLASTN searches against the same database of chloroplast genomes. For genes with low sequence identity, manual annotation was performed, after identifying the position of the start and stop codons, as well as the translated amino acid sequence, using the chloroplast/bacterial genetic code.

Phylogenetic analysis

Complete chloroplast genomes of 10 Poaceae members were used for phylogenetic analysis (*Oryza sativa*, Accession No. NC008155; *Agrostis stolonifera* L., Accession No. NC008591; *Hordeum vulgare* L., Accession No. NC008590; *Triticum aestivum* L., Accession No. NC002762; *Festuca arundinacea* Schreb., Accession No. NC011713; *Lolium perenne* L., Accession No. NC009950.1; *Brachypodium distachyon* L., Accession No. NC011032; *Saccharum offcinarum* L., Accession No. NC006084; *Sorghum bicolor* L., Accession No. NC008602; *Zea mays* L., Accession No. NC001666; Online Supplementary Data). *Dioscorea elephantipes* (Accession No. NC009601) was used as the out-group. Phylogenetic analysis was done by MEGA4 (gap opening penalty: 15; gap extension penalty: 6.66; DNA weight matrix: IUB; transition weight: 0.5; negative matrix: off; delay divergent cutoff: 30%, Tamura et al. 2007). The evolutionary history was inferred using the maximum parsimony and maximum

NIH-PA Author Manuscript

likelihood for phylogenetic analyses (Kolaczkowski and Thornton 2004). In maximum parsimony, the bootstrap consensus tree was inferred from 1000 replicates (Felsenstein 1985). Branches corresponding to partitions reproduced in < 50% bootstrap replicates are collapsed. The values of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (Felsenstein 1985).

Gap region identification

The whole sequences of the *B. oldhamii* and *D. latiflorus* chloroplast genomes were compared with those of the other 10 Poaceae chloroplast genomes. This analysis was done using AlignX, VectorNTI (gap opening penalty: 15; gap extension penalty: 6.66), and gaps of > 100 bp were identified.

Results and discussion

Chloroplast genome sequencing by PCR

In many studies, the first step in sequencing the chloroplast genome has been to establish the plasmid library, BAC library or Fosmid clones. Although these approaches have been successfully used for generating many chloroplast genome sequences, they are not practical for bamboos. No effective method for isolating intact chloroplasts from bamboo has been established for using the rolling circle amplification (RCA) method. The BAC or the Fosmid Library has not yet been constructed for bamboo. In addition, current methods are laborious and have several other limitations. Therefore, we developed a strategy that involved PCR amplification using 443 novel primers (see Online Supplementary Data) to amplify the complete chloroplast genomes of *D. latiflorus* and *B. oldhamii*. As rice and bamboo are closely related, we designed primers based on the coding region of the rice chloroplast genome (primer number 71–414; red primers in Figure 1). The odd numbers were used as forward primers and even numbers as reverse primers (e.g., 71 and 72 as a pair) to perform PCR. The sizes of the PCR products were in the range of 200–2,000 bp. However, there were some gaps. We designed specific primers (primer number 593–1658; M756–M892; blue primers in Figure 1) to fill these gaps. This process helped to reduce the mislinkage during the assembly process, specifically at the junction of single copy and IR regions. The advantage of our method is that the PCR products are sequenced directly without the need for creating libraries, although more recent sequencing methods (454) can directly sequence RCA products without creating libraries.

It is important to pay attention to two points when a PCR-based method is used to sequence the chloroplast genome. First, green leaves are more suitable for deriving total chloroplast DNA because they have more plastid DNA than other organs. Second, the specific primers should be used to reduce the chance of any potential contamination from other organelle DNAs. When intact chloroplasts are isolated for use in the RCA method, contamination of DNA from other genomes is tested in RCA products whereas when total cellular DNA is used as the template, it is not possible to avoid contamination from other genomes. We used leaf DNA as the PCR template and specific primers at a high density for the primary screening. Because of unknown reasons, some primer pairs were not suitable for PCR. For more specific amplification, we used different primer pairs for a specific region. The sequences of the PCR products were then submitted for BLASTN analysis, to evaluate contamination by other organelles. Based on our results, this PCR-based method is a simple (using total DNA as the template) and inexpensive tool (without whole-genome wide sequencer) for sequencing the chloroplast genome. These primers can be used for sequencing the chloroplast genomes of other closely related species and for direct PCR, most specifically members of the Poaceae family.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Chloroplast genome sequencing of bamboo

The complete sizes of the *B. oldhamii* and *D. latiflorus* chloroplast genomes are 139,350 and 139,365 bp, respectively (Figure 2). The genomes include a pair of IRs of 21,790 bp (*B. oldhamii*), and 21,755 bp separated by a small single copy (SSC) region of 12,881 bp (*B. oldhamii*) and 12,854 bp (*D. latiflorus*) and a LSC region of 82,889 bp (*B. oldhamii*) and 83,001 bp (*D. latiflorus*). The *B. oldhamii* and *D. latiflorus* chloroplast genomes contain 112 different genes, of which 19 are duplicated in the IR, giving a total of 131 genes (Figure 2). There are 32 distinct tRNAs, eight of which are duplicated in the IR. Sixteen genes contain one or two introns, and six of these are in tRNAs (protein-encoding genes: *rps16*, *atpF*, *ycf3*, *petD*, *rpl16*, *rpl2*, *rps12*, *ndhB*, *ndhA*, *petB*; tRNA: UUU, UAC, GAU, UGC, UCC, UAA). Coding regions make up 53.4% of the *B. oldhamii* and *D. latiflorus* chloroplast genomes (44.8% protein-coding genes and 8.6% RNA genes) and non-coding regions, 46.6% (contains intergenic spacer regions and introns). The overall GC and AT content of the chloroplast genome is 39.0% and 61.0%, respectively, for *B. oldhamii* and 38.9% and 61.1%, respectively, for *D. latiflorus*. However, the IR regions of both the *B. oldhamii* and *D. latiflorus* chloroplast genomes have an AT content of 55%, which is similar to that of the chloroplast genomes of other Poaceae and *F. esculentum* ssp. *ancestrale* (Logacheva et al. 2008).

Gene content and gene order

The gene content and gene order of the *B. oldhamii* and *D. latiflorus* chloroplast genomes are identical to those of other sequenced grass chloroplast genomes, including the loss of introns in *clpP* and *rpoC1* and loss of *accD*, *ycf1* and *ycf2* genes. Like other grass chloroplast genomes, the IR in *B. oldhamii* and *D. latiflorus* has expanded to include *rps19*. Furthermore, between two of the genomes with the IR of *B. oldhamii* and *D. latiflorus*, the IR has expanded to duplicate a portion of *ndhH*, a feature that is similar to the results found in other Poaceae species (Ogihara et al. 2002, Saski et al. 2007). This expansion is 198 bp (66 amino acids) in *B. oldhamii* and *D. latiflorus* (Online Supplementary Data). These data indicate that the expansion of *ndhH* is a specific characteristic of BEP clade grasses but not PACCAD clade members.

Gaps in Poaceae chloroplast genomes

Insertions and deletions (indels) of nucleotides within several coding sequences have been found in chloroplast genomes. For example, a 6 bp sequence is uniquely present within *matK* of *Sorghum bicolor*, but absent in the rest of the grasses, including the two bamboo species analyzed here (Saski et al. 2007). Also, 6 bp in the *ndhK* gene was deleted in bamboo chloroplast genomes. This indicates that this deletion is BEP clade-specific (Saski et al. 2007).

Insertions and deletions have also been found in the chloroplast genomes of grasses. Bortiri et al. (2008) reported a 2.1-kb deletion in rice and wheat. A smaller deletion (1.4 kb) appears in other BEP clade members (Figure 3). This deletion is BEP-specific. A comparison of our two bamboo chloroplast genomes with those of other Poaceae chloroplast genomes revealed 23 deletions (Table 1). There were seven indels in 1–20,000. Relatively speaking, members of the Pooideae have more indels. One is PACCD-clade specific (gap 18) and three Pooideae deletions can be identified (gaps 9, 14 and 19). Two more deletions only appear in *T. aestivum* (gaps 4 and 11). One specific deletion in rice (gap 21), *B. distachyon* (gap 8) and *F. arundinacea* (gap 16) were also identified. These gaps can be used as indicators of evolution and phylogeny.

Phylogenetic analysis

To determine the relative position of *B. oldhamii* and *D. latiflorus*, we adopted the method of Saski et al. (2007) and established a phylogenetic tree of 38 taxa based on 61 plastid protein-

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

coding genes using maximum parsimony (data not shown). Our results are in full agreement with those of Saski et al. (2007). However, the Bambusoideae is a sister subfamily to the Erhartoideae and Pooideae with weak-to-moderate bootstrap support. Because this method is based on only 61 of the protein-coding region sequences, which are highly conserved in Poaeceae, it would be preferable to use the complete chloroplast genome to analyze closely related species, as demonstrated by Bortiri et al. (2008). Based on a comparison of the entire chloroplast genomes, we found that bamboos are related to Erhartoideae (Figure 4). We therefore suggest using a complete chloroplast genome comparison to analyze the relationship among Poaceae.

Conclusions

The complete chloroplast genome sequences of two economically important woody bamboos, *B. oldhamii* and *D. latiflorus*, have been completed using an inexpensive and simple PCR-based method. The primers designed and the methods developed can be used for sequencing other chloroplast genomes, specifically those in the Poaceae family. The identified sequences of the *B. oldhamii* and *D. latiflorus* chloroplast genomes will be useful for bamboo biotechnology research and in determining the mechanism of mutant regeneration in tissue culture. There are 20 gaps of > 100 bp in our chloroplast genome sequences; these were identified based on a comparison with the chloroplast genomes of other Poaceae species. The gap polymorphisms are in agreement with the relationships among the major clades of grasses. We have identified PACCAD-clade specific sequences and a Pooideae-specific gap. In terms of phylogenetic studies of the Poaceae family, these two chloroplast genome sequences are the first to be reported for Bambusoideae. The complete genome analysis adds to the knowledge on the phylogeny of the Poaceae, specifically in the BEP clade. Because *Bambusa* and *Dendrocalamus* are closely related and as it is difficult to obtain reproductive tissues, a comparison of the entire chloroplast genomes will be a useful method to study the phylogeny of the members of the BEP clade at the genus, species level, or both.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study is supported by the National Science Council and Academia Sinica, Taipei, Taiwan. The authors are grateful to W.C. Chang for providing facility support and Dr N.P. Samson in the Daniell lab for drawing the genome map.

References

Arlen PA, Falconer R, Cherukumilli S, Cole A, Cole AM, Oishi KK, Daniell H. Field production and functional evaluation of chloroplast-derived interferon-alpha2b. Plant Biotech J 2007;5:511–525.

Arlen PA, Singleton M, Adamovicz JJ, Ding Y, Davoodi-Semiromi A, Daniell H. Effective plague vaccination via oral delivery of plant cells expressing F1-V antigens in chloroplasts. Infect Immun 2008;76:3640–3650. [PubMed: 18505806]

Bortiri E, Coleman-Derr D, Lazo GR, Anderson OD, Gu YQ. The complete chloroplast genome sequence of *Brachypodium distachyon*: sequence comparison and phylogenetic analysis of eight grass plastomes. BMC Research Notes 2008;1:61. [PubMed: 18710514]

Daniell H. Transgene containment via maternal inheritance. Proc Natl Acad Sci USA 2007;104:6879–6881. [PubMed: 17440039]

Daniell H, Datta R, Varma S, Gray S, Lee SB. Containment of herbicide resistance through genetic engineering of the chloroplast genome. Nat Biotechnol 1998;16:345–348. [PubMed: 9555724]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Daniell H, Lee SB, Grevich J, Saski C, Guda C, Tomkins J, Jansen RK. Complete chloroplast genome sequences of *Solanum bulbocastanum, Solanum lycopersicum* and comparative analyses with other *Solanaceae genomes*. Theor Appl Genet 2006;112:1503–1518. [PubMed: 16575560]

Daniell H, Wurdack KJ, Kanagaraj A, Lee SB, Saski C, Jansen RK. The complete nucleotide sequence of the cassava (*Manihot esculenta*) chloroplast genome and the evolution of *atpF* in Malphigiales: RNA editing and multiple losses of a group II intron. Theor Appl Genet 2008;116:723–737. [PubMed: 18214421]

Day A, Ellis TH. Chloroplast DNA deletions associated with wheat plants regenerated from pollen: possible basis for maternal inheritance of chloroplasts. Cell 1984;39:359–368. [PubMed: 6094013]

DeCosa B, Moar W, Lee SB, Miller M, Daniell H. Overexpression of the *Bt Cry2Aa2* operon in chloroplasts leads to formation of insecticidal crystals. Nat Biotechnol 2001;19:71–74. [PubMed: 11135556]

DeGray G, Rajasekaran K, Smith F, Sanford J, Daniell H. Expression of an antimicrobial peptide via the chloroplast genome to control phytopathogenic bacteria and fungi. Plant Physiol 2001;127:852–862. [PubMed: 11706168]

Demesure B, Sodzi N, Petit RJ. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol 1995;4:129–131. [PubMed: 7711952]

Dhingra A, Folta KM. ASAP: amplification, sequencing, and annotation of plastomes. BMC Genomics 2005;6:176. [PubMed: 16336644]

Dumolin-Lapegue S, Pemonge MH, Petit RJ. An enlarged set of consensus primers for the study of organelle DNA in plants. Mol Ecol 1997;6:393–398. [PubMed: 9131816]

Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 1985;39:783–791.

Grass Phylogeny Working Group. Phylogeny and subfamilial classification of the grasses (Poaceae). Annals Missouri Bot Gard 2001;88:373–457.

Grevich JJ, Daniell H. Chloroplast genetic engineering: recent advances and future perspectives. Crit Rev Plant Sci 2005;24:83–107.

Grivet D, Heinze B, Vendramin GG, Petit RJ. Genome walking with consensus primers: application to the large single copy region of chloroplast DNA. Mol Ecol Notes 2001;1:345–349.

Heinze B. A database of PCR primers for the chloroplast genomes of higher plants. Plant Methods 2007;3:4–11. [PubMed: 17326828]

Hsu YH, Annamalai AP, Lin CS, Chen YY, Chang WC, Lin NS. A sensitive method for detecting bamboo mosaic virus (BaMV) and establishment of BaMV-free meristem tip cultures. Plant Path 2000;49:101–107.

Jansen RK, Raubeson LA, Boore JL, et al. Methods for obtaining and analyzing whole chloroplast genome sequences. Meth Enzymol 2005;395:348–384. [PubMed: 15865976]

Jansen RK, Kaittanis C, Saski C, Lee SB, Tomkins J, Alverson AJ, Daniell H. Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. BMC Evol Biol 2006;6:32–46. [PubMed: 16603088]

Kode V, Mudd EA, Iamtham S, Day A. Isolation of precise plastid deletion mutants by homology-based excision: a resource for site-directed mutagenesis, multi-gene changes and high-throughput plastid transformation. Plant J 2006;46:901–909. [PubMed: 16709203]

Kolaczkowski B, Thornton JW. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. Nature 2004;431:980–984. [PubMed: 15496922]

Koya V, Moayeri M, Leppla SH, Daniell H. Plant-based vaccine: mice immunized with chloroplast-derived anthrax protective antigen survive anthrax lethal toxin challenge. Infect Immun 2005;73:8266–8274. [PubMed: 16299323]

Kumar S, Dhingra A, Daniell H. Chloroplast-expressed *betaine aldehyde dehydrogenase* gene in carrot cultured cells, roots and leaves confers enhanced salt tolerance. Plant Physiol 2004;136:2843–2854. [PubMed: 15347789]

Lee SB, Kwon HK, Kwon SJ, Park SC, Jeong MJ, Han SE, Byun MO, Daniell H. Accumulation of trehalose within transgenic chloroplasts confers drought tolerance. Mol Breed 2003;11:1–13.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Lee SM, Kang K, Chung H, Yoo SH, Xu XM, Lee SB, Cheong JJ, Daniell H, Kim M. Plastid transformation in the monocotyledonous cereal crop, rice (*Oryza sativa*) and transmission of transgenes to their progeny. Mol Cells 2006a;21:401–410. [PubMed: 16819304]

Lee SB, Kaittanis C, Jansen RK, Hostetler JB, Tallon LJ, Town CD, Daniell H. The complete chloroplast genome sequence of *Gossypium hirsutum*: organization and phylogenetic relationships to other angiosperms. BMC Genomics 2006b;7:61–77. [PubMed: 16553962]

Leelavathi S V, Reddy S. Chloroplast expression of His-tagged GUS-fusions: a general strategy to overproduce and purify foreign proteins using transplastomic plants as bioreactors. Mol Breed 2003;11:49–58.

Lin CS, Chang WC. Micropropagation of *Bambusa edulis* through nodal explants of field-grown culms and flowering of regenerated plantlets. Plant Cell Rep 1998;17:617–620.

Lin CS, Liang CJ, Hsaio HW, Lin MJ, Chang WC. Flowering of the green and albino regenerates obtained via somatic embryogenesis in inflorescence segments of *Dendrocalamus latiflorus*. New Forests 2007;34:177–186.

Lin CS, Liu NT, Liao DC, et al. Differential protein expression of two photosystem II subunits, PsbO and PsbP in an albino mutant of *Bambusa edulis* Munro with chloroplast DNA aberration. J Amer Soc Hort Sci 2008;133:270–277.

Liu NT, Jane WN, Tsay HS, Wu H, Chang WC, Lin CS. Chloroplast genome aberration in micropropagation derived albino *Bambusa edulis* mutants, *ab1* and *ab2*. Plant Cell Tiss Org Cult 2007;88:147–156.

Logacheva MD, Samigullin TH, Dhingra A, Penin AA. Comparative chloroplast genomics and phylogenetics of *Fagopyrum esculentum* ssp. ancestrale – a wild ancestor of cultivated buckwheat. BMC Plant Biol 2008;8:59. [PubMed: 18492277]

Martin W, Stoebe B, Goremykin V, Hapsmann S, Hasegawa M, Kowallik KV. Gene transfer to the nucleus and the evolution of chloroplasts. Nature 1998;393:162–165. [PubMed: 11560168]

McBride KE, Svab Z, Schaaf DJ, Hogan PS, Stalker DM, Maliga P. Amplification of a chimeric *Bacillus* gene in chloroplasts leads to an extraordinary level of an insecticidal protein in tobacco. BioTechn 1995;13:362–365.

Ogihara Y, Isono K, Kojima T, et al. Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. Mol Genot Genomics 2002;266:740–746.

Palmer JD. Chloroplast DNA and molecular phylogeny. Bioessays 1985;2:263–267.

Quesada-Vargas T, Ruiz ON, Daniell H. Characterization of heterologous multigene operons in transgenic chloroplast. Transcription, processing and translation. Plant Physiol 2005;138:1746–1762. [PubMed: 15980187]

Ruhlman T, Lee SB, Jansen RK, Hostetler JB, Tallon LJ, Town CD, Daniell H. Complete plastid genome sequence of *Daucus carota*: implications for biotechnology and phylogeny of angiosperms. BMC Genomics 2006;7:224. [PubMed: 16948847]

Ruiz ON, Hussein HS, Terry N, Daniell H. Phytoremediation of organomercurial compounds via chloroplast genetic engineering. Plant Physiol 2003;132:1344–1352. [PubMed: 12857816]

Samson N, Bausher MG, Lee SB, Jansen RK, Daniell H. The complete nucleotide sequence of the coffee (*Coffea arabica* L.) chloroplast genome: organization and implications for biotechnology and phylogenetic relationships among angiosperms. Plant Biotech J 2007;5:339–353.

Saski C, Lee SB, Daniell H, Wood T, Tomkins J, Kim HG, Jansen R. Complete chloroplast genome sequence of *Glycine max* and comparative analyses with other legume genomics. Plant Mol Biol 2005;59:309–322. [PubMed: 16247559]

Saski C, Lee SB, Fjellheim S, Guda C, Jansen RK, Luo H, Tomkins J, Rognli OA, Daniell H, Clarke JL. Complete chloroplast genome sequences of *Hordeum vulgare, Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with other grass genomes. Theor Appl Genet 2007;115:571–590. [PubMed: 17534593]

Sheu JJ, Yu TS, Tong WF, Yu SM. Carbohydrate starvation stimulates differential expression of rice α-amylase genes that is modulated through complicated transcriptional and postranscriptional process. J Biol Chem 1996;27:26998–27004. [PubMed: 8900187]

Stapleton, CMA. Morphology of woody bamboos. In: Chapman, GP., editor. The Bamboos. Academic Press; 1997. p. 251-267.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Staub JM, Garcia B, Graves J, et al. High-yield production of a human therapeutic protein in tobacco chloroplasts. Nat Biotechnol 2000;18:333–338. [PubMed: 10700152]

Tamura K, Dudley J, Nei M, Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software Version 4.0. Mol Biol Evol 2007;24:1596–1599. [PubMed: 17488738]

Tang J, Xia H, Cao M, et al. A comparison of rice chloroplast genome. Plant Physiol 2004;135:412–420. [PubMed: 15122023]

Tregoning JS, Clare S, Bowe F, Edwards L, Fairweather N, Qazi O, Nixon PJ, Maliga P, Dougan G, Hussell T. Protection against tetanus toxin using a plant-based vaccine. Eur J Immunol 2005;35:1320–1326. [PubMed: 15739166]

Verma D, Daniell H. Chloroplast vector systems for biotechnology applications. Plant Physiol 2007;145:1129–1143. [PubMed: 18056863]

Verma D, Samson NP, Koya V, Daniell H. A protocol for expression of foreign genes in chloroplasts. Nat Protocols 2008;3:739–758.

Wyman SK, Jansen RK, Boore JL. Automatic annotation of organellar genomes with DOGMA. Bioinformatics 2004;20:3252–3255. [PubMed: 15180927]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

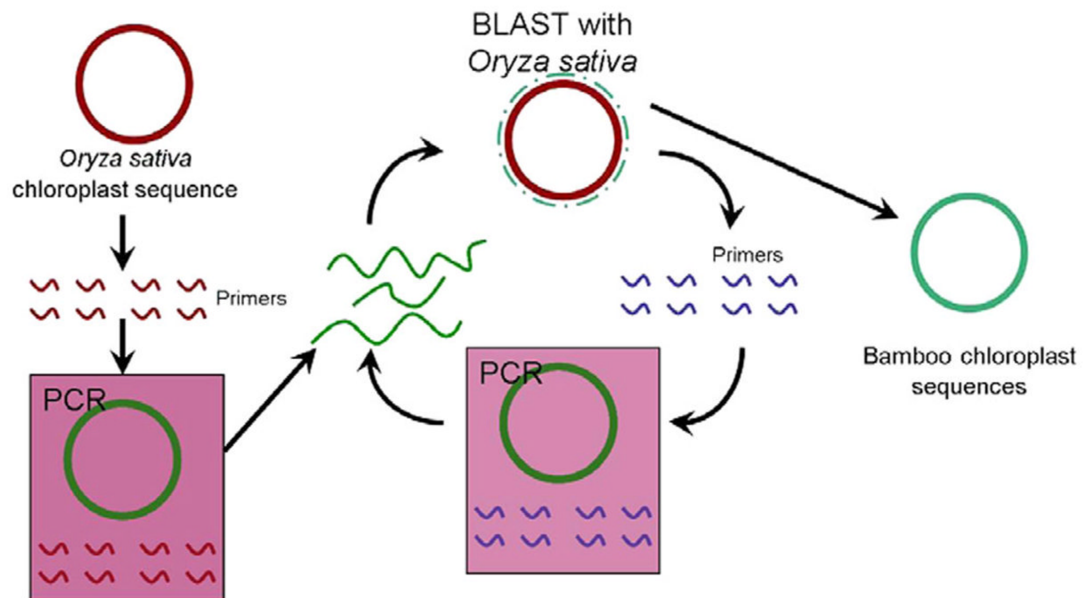

Figure 1.
Strategy of bamboo chloroplast genome sequencing. Based on the rice (*Oryza sativa* L.) sequences, the primers are designed (red primers). Using the total DNA of *B. oldhamii* and *D. latiflorus*, PCR was performed. The PCR products were sequenced and BLAST analysis was performed with the rice chloroplast genome. The gap sequences were filled by designing new primers (blue primers) for further PCR. Red circle: rice chloroplast genome; green circle: bamboo chloroplast genome; purple box: PCR.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

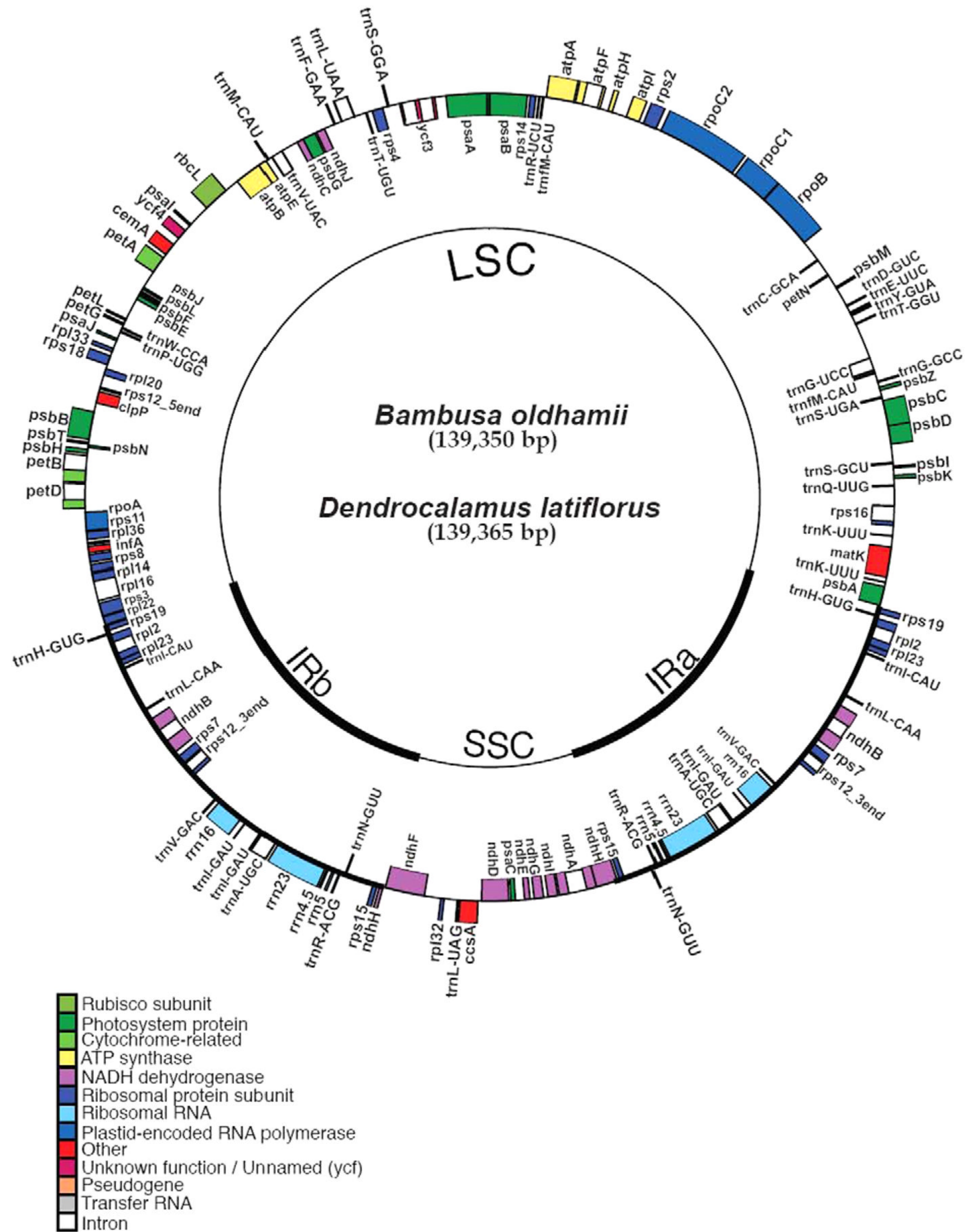

Figure 2.

Gene map of *B. oldhamii* and *D. latiflorus* chloroplast genomes. The thick lines indicate the extent of the IRa and IRb, which separate the genome into SSC and LSC regions. Genes on the outside of the map are transcribed in the clockwise direction and genes on the inside of the map are transcribed in the counterclockwise direction.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

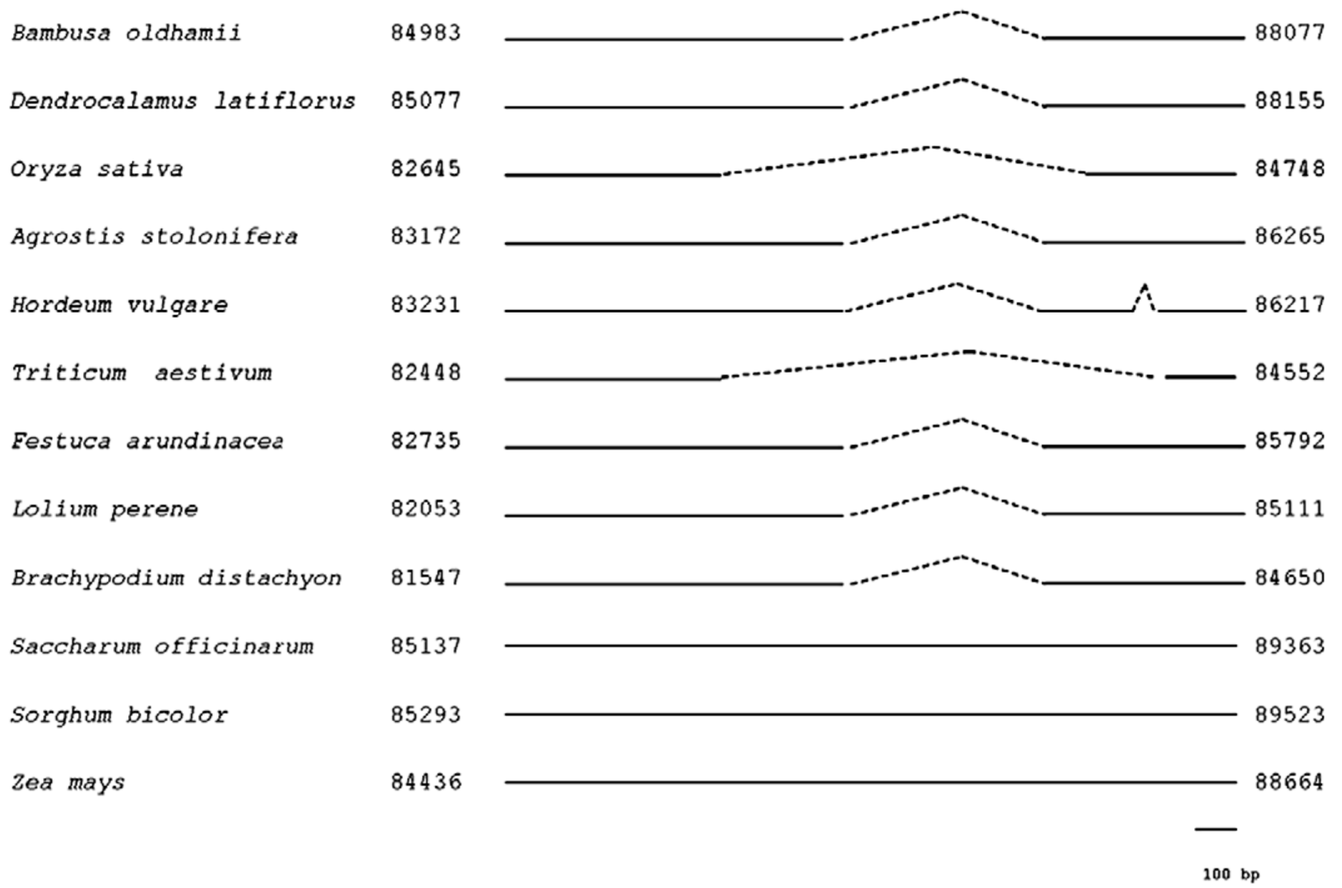

Figure 3.
Deletions in the IRa region. IRa regions of *B. oldhamii* and *D. latiflorus* chloroplast genomes are similar to *A. stolonifera* and *H. vulgare*, which have a 1.4-kb deletion. This phenomenon also appears in the IRa of the *B. distachyon* chloroplast genome (Bortiri et al. 2008). Details of the gap number 18 are given in Table 1.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

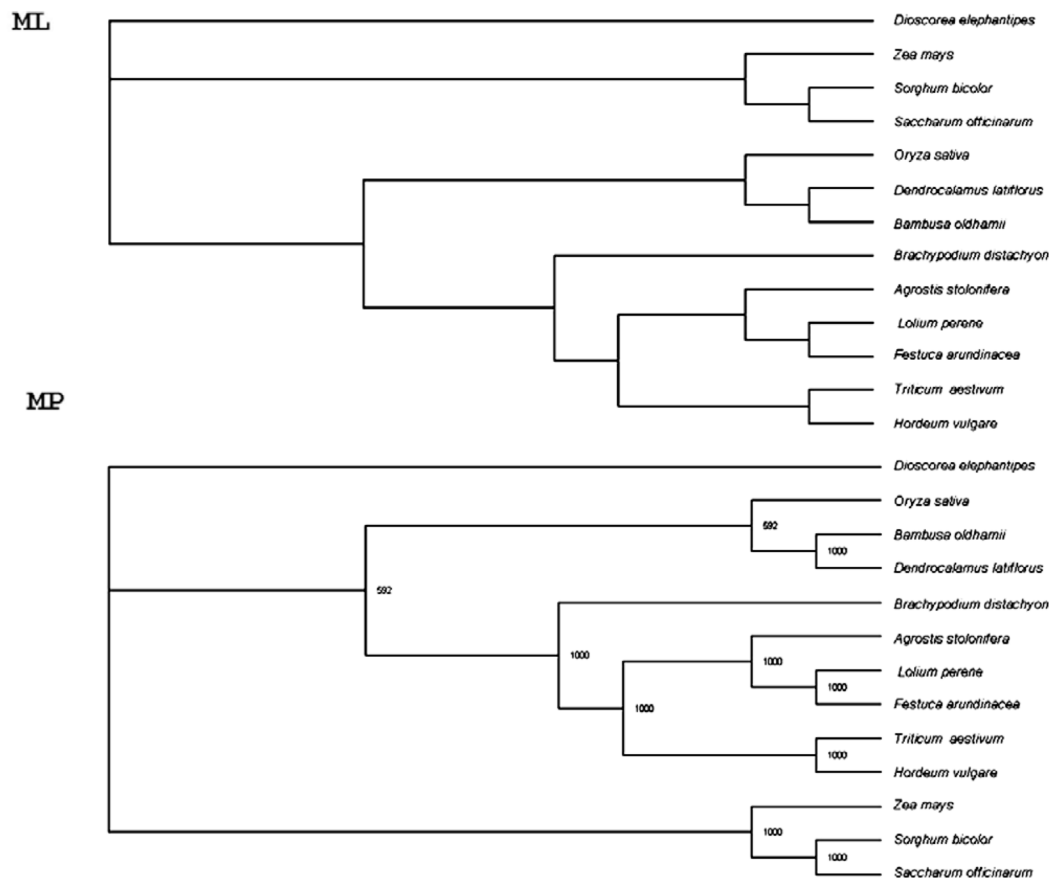

Figure 4.

The tree-explorer displaying evolutionary relationships of 12 Poaceae species by using the chloroplast genome. Phylogenetic analyses were conducted in MEGA4 (ML, maximum likelihood; MP, maximum parsimony; Tamura et al. 2007).

Table 1

Summary of the gaps in the Poaceae. (A) Gap size. White triangle: the gap size is < 500 bp; grey triangle: the gap size is between 500 and 1000 bp and black triangle: the gap size is > 1000 bp. (B) Gap analysis of the *B. oldhamii* chloroplast genome.

A.

Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
B. oldhamii	△																△	▲(grey)					▲
D. latiflorus	△																	▲(grey)					▲
O. sativa	△	△	△		▲(grey)		▲(grey)											▲			△	△	▲
A. stolonifera	▲(grey)		▲(grey)		△	▲(grey)			△	△			△	△				▲(grey)	△				▲
H. vulgare	▲(grey)		▲(grey)		△		△		△	△		△	△	△	△			▲(grey)	△				▲
T. Aestivum	△	△	▲(grey)	△	△	△			△	△	△		▲(grey)	△	△			▲	△				▲
F. arundinacea	▲(grey)		▲(grey)		△	▲(grey)			△			△	▲(grey)	△		△	△	▲(grey)	△			△	▲
L. perene	▲(grey)		▲(grey)		△	▲(grey)			△			△	△	△			△	▲(grey)	△			△	▲
B. distachycn	▲(grey)	△	▲(grey)		△			△	△				▲(grey)	△				▲(grey)	△			△	▲
S. officinarum										△			▲(grey)							△			▲
S. bicolor										△			▲(grey)							△		△	
Z. mays	△	△								△			▲(grey)							△		△	

△:<0.5K ▲(grey):0.5–1K ▲:>1K

B.

Gap	1	2	3	4	5	6	7	8	9	10
Start	5945	13816	14592	15951	16947	18319	18867	33354	33532	47960
End	6530	14388	16014	16385	18029	18939	19900	33565	34136	49209
Gap	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
Start	49576	52401	58078	62225	64720	67839	80453	84983	96735	104075
End	50261	52977	59292	63679	65068	68039	80782	88077	97250	105870
Gap	**21**	**22**	**23**							
Start	107213	108256	134794							
End	107640	108892	671							

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript