

The Complete Chloroplast Genome of Pearl Millet (*Pennisetum glaucum* (L.) R. Br.) and Comparative Analysis within the Family Poaceae

S. RAVEENDAR¹, G.A. LEE¹, K.J. LEE¹, M.J. SHIN¹, J.R. LEE¹, S.Y. LEE¹, G.T. CHO¹,
K.H. MA^{1*} and J.W. CHUNG^{2*}

¹National Agrobiodiversity Center, National Institute of Agricultural Science, Rural Development Administration, Jeonju 54874, Republic of Korea

²Department of Industrial Plant Science and Technology, Chungbuk National University, Cheongju, Chungbuk 28644, Republic of Korea

(Received 29 May 2018; Accepted 7 September 2018;

Communicated by M. Molnár-Láng)

The complete chloroplast (cp) genome sequence of Pearl millet (*Pennisetum glaucum* [L.] R. Br.), an important grain and forage crop in the family Poaceae, is reported in this study. The complete cp genome sequence of *P. glaucum* is 138,172 bp in length with 38.6% overall GC content and exhibits a typical quadripartite structure comprising one pair of inverted repeat (IR) regions (22,275 bp) separated by a small single-copy (SSC) region (12,409 bp) and a large single-copy (LSC) region (81,213). The *P. glaucum* cp genome encodes 110 unique genes, 76 of which are protein-coding genes, 4 ribosomal RNA (rRNA) genes, 30 transfer RNA (tRNA) genes and 18 duplicated genes in the IR region. Nine genes contain one or two introns. Whole genome alignments of cp genome were performed for genome-wide comparison. Locally collinear blocks (LCBs) identified among the cp genomes showed that they were well conserved with respect to gene organization and order. This newly determined cp genome sequence of *P. glaucum* will provide valuable information for the future breeding programs of valuable cereal crops in the family Poaceae.

Keywords: chloroplast, Pearl millet, phylogenetic analysis, quadripartite structure

Introduction

Pearl millet [*Pennisetum glaucum* (L.) R. Br.] is a widely cultivated C₄ cereal and mainly grown under dryland, rainfed and irrigated conditions in the world's arid tropical and subtropical especially Africa and India (ICRISAT 1996). The great significance of this crop has the ability to grow rapidly and produce seed in a short season is evidenced as an important staple food grain, and source of feed and fodder for livestock, in hot, dry marginal agricultural production environments of arid and semiarid regions (Burton 1983). Also, several other economic importance and utilization of pearl millet have been reported (Poncet et al. 1998).

*Corresponding authors; E-mails: jwchung73@chungbuk.ac.kr, khma@korea.kr

Pearl millet pose a genetically closer relationship with several biofuel grasses, such as foxtail millet (*Setaria italica*), napiergrass (*P. purpureum*), switchgrass (*Panicum virgatum*) and so on, and due to its highly out-crossing breeding behavior, a large number of diversity at both phenotypic (Poncet et al. 1998) and genotypic levels (Govindaraj et al. 2009) is found within and among pearl millet cultivars. Efficient and systematic exploitation of these diversity and population structure is important for germplasm conservation and cultivar development (Ramakrishnan et al. 2016). Several studies have been conducted to understand the genetic diversity of pearl millet germplasm using molecular markers. The first molecular marker-based genetic linkage map using restriction fragment length polymorphism (RFLP) markers is a major milestone in pearl millet (Liu et al. 1994).

Recently, the next generation sequencing (NGS) technologies have been used in many crops to rapidly identify genomic variation and develop genetic maps (Huang et al. 2012; Rajaram et al. 2013; Wu et al. 2015). The major advantages of its small genome size, rich genetic diversity, Pearl millet promises to be an important model crop for comparative genomics and functional gene studies (Hu et al. 2015). The genetic resources for nuclear and chloroplast genomes are available for several grasses and related species in the Poaceae family. However, currently available genetic resources for the Pearl millet are very limited. Therefore, we characterized the complete chloroplast genome sequence of a grass Pearl millet for the first time, aiming at providing useful information for further genetic and phylogenetic studies within Poaceae family.

Materials and Methods

Plant and DNA sample preparation

Pearl millet seeds (Accession No. IT144339) were collected from the National Agrobiodiversity Center, Rural Development Administration (RDA), Republic of Korea. Seeds were germinated and grown in a greenhouse, and fresh leaves were collected from 40-day-old seedlings, and DNA was extracted using the modified cetyltrimethylammonium bromide (CTAB) method (Allen et al. 2006). The entire process of DNA sample preparation for the whole genome shotgun sequencing was performed according to the manufacturer's instructions of the Illumina platform (San Diego, CA, USA).

Chloroplast genome sequencing, assembly, and validation

An Illumina paired-end (PE) library was sequenced by an Illumina HiSeq2000 genome analyzer platform (Illumina, San Diego, CA, USA) installed in Macrogen (Macrogen, Seoul, Korea) and cp genome was obtained by *de novo* assembly of the low coverage whole genome sequence via a bioinformatics pipeline (<http://phyzen.com>). Low quality sequences (Phred score < 20) were trimmed and the remaining high quality sequences were assembled in to contigs using a CLC genome assembler beta 4.06 (CLC Inc., Rarhus, Denmark) with parameters of a minimum of 200–600 bp autonomously controlled over-

lap size at Phyzen Inc. (Seoul, South Korea). The principal contigs representing the chloroplast genome were obtained from the total assembled contigs using MUMmer (Kurtz et al. 2004) with the cp genome sequence of *Avena sativa* (KM974733) as a reference sequence (Saarela et al. 2015). The representative cp contigs were arranged by performing a BLAST (ver. 2.2.31) analysis with the reference sequence and ambiguous nucleotides or gaps were corrected manually to build the complete cp genome.

Chloroplast genome annotation and repeat sequence analysis

Gene annotation was conducted using the web-based program Dual OrganellarGenoMe Annotator (DOGMA) (Wyman et al. 2004) and manual editing through comparison with the reported cp genome sequence of *Avena sativa* (KM974733). The circle map of the *P. glaucum* cp genome was obtained using OrganellarGenomeDRAW software (OGDRAW, <http://ogdraw.mpimp-golm.mpg.de>) (Lohse et al. 2013). Chloroplast SSR (cpSSR) markers present in the cp genome was identified with MicroSatellite identification tool (MISA) software (<http://pgrc.ipk-gatersleben.de/misa/>). MISA has been applied for SSR identification in many species including Eucalyptus (Ceresini et al. 2005) and barley (Thiel et al. 2003).

Codon usage

A total of 83 protein coding genes in the cp genome of *P. glaucum* were concatenated to calculate the synonymous codon usage analysis. Number of codons (Nc) and relative synonymous codon usage (RSCU) parameters were calculated using the program CodonW1.4.2 (Available online: <http://downloads.fyxm.net/CodonW-76666.html>). Nc used in a gene is a simple measure of codon bias (Wright 1990) in which the Nc values can vary from 20 (extreme bias) to 61 (totally unbiased). RSCU is the observed frequency of a codon divided by the expected frequency (Sharp et al. 1986). If all synonymous codons coding the same amino acid were used equally, RSCU values were close to 1.0, indicating a lack of bias. The amino acid (AA) frequency was also calculated and expressed by the percentage of the codons encoding the same amino acid divided by the total codons.

Comparative analysis with common buckwheat

The cp genome sequences representing Poaceae family related species were downloaded from the NCBI database. The progressive alignment method, implemented in the MAUVE v.2.3.1 package was used with the default parameters to identify locally collinear blocks (LCBs) among the cp genomes (Darling et al. 2004). The locally collinear blocks (LCBs) identified by the Mauve alignment are indicated by color. Inferring the phylogenetic relationship among the species, a new multiple sequence alignment (MSA) was performed by adding 18 Poaceae family cp genome sequences. The resulting alignment was imported in MEGA 6.0 for inferring maximum parsimony (MP) analysis, setting 1000 bootstrap replications, Tree-Bisection-Regrafting (TBR) algorithm and 10 random addition repli-

cates. All positions containing gaps and missing data were eliminated. Phylogenetic tree was generated by a maximum likelihood (ML) analysis using MEGA 6.0 (<http://www.megasoftware.net/>) with 1000 bootstrap replicates (Tamura et al. 2013).

Results

We sequenced the cp genome of Pearl millet (*P. glaucum*) using the Illumina genome analyzer platform. Illumina paired-end (2×300 bp) sequencing produced a total of 6,830,884 paired-end reads, with an average fragment length of 256 bp, which were then analyzed to generate 2,049,481,261 bp of sequence. Low quality reads (Q20) were filtered out, and the remaining high quality reads were mapped to the reference cp genome of *Avena sativa* (GenBank accession KM974733), which contains 81,090 aligned reads with an average coverage of $129\times$ on the cp genome. The cp genome of *P. glaucum* has been deposited to NCBI GenBank with Synonym species name *Cenchrus americanus* under accession number KX756179.

The total length of the chloroplast genome is 138,172 bp, with 38.6% overall GC content. A pair of IRs (inverted repeats) of 22,275 bp was separated by a small single copy (SSC) region of 12,409 bp and a large single copy (LSC) region of 81,213 bp (Fig. 1). The chloroplast genome harbors 110 known genes, including 76 protein-coding genes, 4 rRNA genes, and 30 tRNA genes. A total of 18 genes were duplicated in the inverted repeat regions, eight genes and two tRNA genes contain one intron, while *ycf3* have two introns (Table S1*). Trans-splicing of *rps12* gene is observed in the cp genome, as is the case with other cp genomes in angiosperm. Moreover, the cp genome has also lost *accD*, *ycf1*, and *ycf2* genes.

A total of 19,789 codons were analyzed with the 83 coding genes. The amino acid (AA) frequencies of the *P. glaucum* cp genome were further computed. Of these codons, 2150 (10.90%) encode leucine, which was the most frequency used amino acid in the *P. glaucum* cp genome and cysteine was only encoded by 220 (1.11%) codons which was the least used one (Table S2). The effective number of codons (Nc) for the 83 coding genes from the *P. glaucum* is present in Table S2. The Nc values for each coding genes are shown in Table S3. The results indicated that the Nc values were greater than 40, which suggested a weak gene codon bias in the *P. glaucum* cp genome. The overall codon and relative synonymous codon usage (RSCU) analysis showed that there are 29 codons with RSCU values of >1 , which indicates codon usage bias in the *P. glaucum* cp genes. Interestingly, out of the above 29 codons, twenty-eight were A or U-ending codons.

The occurrence, type, and distribution of chloroplast SSRs were analyzed in the *P. glaucum* cp genome with the MISA perlscript. The analysis produced a total of 77 cp microsatellites in the *P. glaucum* cp genome (Table S4). Out of the 77 cpSSRs, 41 repeats were found to be tri-nucleotides (53.2%), 22 repeats were di-nucleotides (28.6%) and 14 repeats were tetra-nucleotides (18%). Meanwhile, 54.5% of repeats were located in coding intergenic region and 45.5% of repeats were in the intergenic region. These molecular

*Further details about the Electronic Supplementary Material (ESM) can be found at the end of the article.

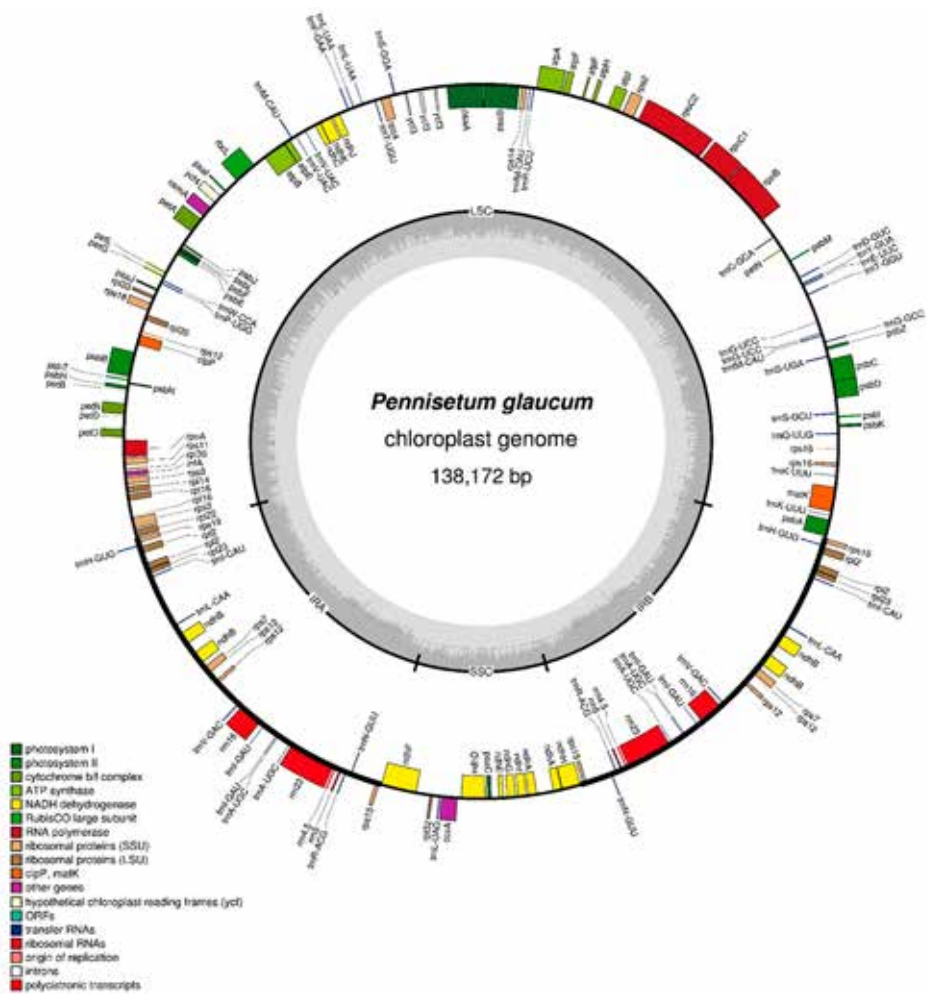


Figure 1. Complete genome map of *P. glaucum* chloroplast. Genes drawn inside the circle are transcribed clockwise, while those outside are counterclockwise and marked with two arrows. Differential functional gene groups are color-coded. The GC content variation is shown in the middle circle

markers will facilitate the studies of genetic diversity, population genetic structure, and sustainable conservation for *P. glaucum*.

Whole-genome alignment of nine *Poaceae* species was performed in Mauve V 2.3.1 and compared the LSC/IRb/SSC/IRA region of the cp genomes (Fig. S1). The locally collinear blocks (LCBs) identified by the Mauve alignment showed that they were well conserved with respect to gene organization and many gene clusters were shared specifically among the cp genomes. Similarly, the LCBs were well conserved with respect to both gene identity and order among the *Poaceae* cp genomes.

To identify the evolutionary position of *P. glaucum* within Poaceae, an improved resolution of phylogenetic relationships was achieved by using these whole cp genome sequences of 18 Poaceae species. The phylogenetic analysis indicated that Pearl millet clustered with *Setaria viridis* (common name: green foxtail) and *Setaria italica* (common name: foxtail millet) with 100% bootstrap supports. Both *S. italica* and *S. viridis* are members of the tribe Paniceae that is the sister to the Andropogoneae. The maximum likelihood (ML) bootstrap values were high with 100% bootstrap support in almost all nodes except the sub-cluster of *Brachypodium distachyon* (Fig. S2) which generally comes under the BOP grasses which use the C₃ photosynthetic pathway.

Discussion

Pearl millet is an economically important species used as a staple food grain and source of feed and fodder for livestock in Africa and South Asia (ICRISAT 1996). The assembled *P. glaucum* cp genome showed a high similarity to the reference sequence, *Avena sativa* (KM974733) and extended its size to 138,172 nucleotides long. As shown in Figure 1, the *P. glaucum* cp genome has a typical quadripartite structure composed of the LSC, SSC and a pair of IR regions, indicating that the assembled *P. glaucum* cp genome sequence represented full coverage with no abnormalities. Several chloroplast genome sequences have been reported from the Poaceae family (Cao et al. 2017; Huang et al. 2017; Liu et al. 2017). The cp genome of *P. glaucum* is highly conserved, with identical gene content and gene order and genomic structure comprising four parts as reported as angiosperm cp genomes. The average GC content was 38.6%, which is almost identical with each other among the other reported Poaceae cp genomes (Cao et al. 2017; Huang et al. 2017; Liu et al. 2017).

Various biological factors and different approaches have been demonstrated for codon usage bias determination (Cannarozzi and Schneider 2012). However, in seed plants genomic nucleotide mutation bias considered as a major factors for codon usage variation (Zhou et al. 2008). In a relevant study (Zhang et al. 2012), suggesting that mutational bias played a major role in the codon biology of the Poaceae chloroplast genome. In the present study, our results indicated that most of Nc values were greater than 40 and the RSCU values of >1 suggested a weak gene codon bias, which may due to high A or U-ending codons in the *P. glaucum* cp genome. The similar codon usage bias for A or U-ending was also found in other plants (Zhang et al. 2012; Zhou et al. 2008). In general, the stop codon prefers to use UAA more often and the synonymous codon usage can be easily examined by the RSCU values. We also found that the start codons of the *rpl2* and *rps19* genes are likely to convert to ACG and GTG via RNA editing during translation, as reported in other species (Gao et al. 2009; Kahlau et al. 2006).

Chloroplast microsatellites (cpSSRs) are considered suitable molecular markers because of their locus-specific co-dominance and high polymorphic nature which makes them as useful for phylogenetic relationships analysis in land plants (Olmstead and Palmer 1994). Moreover, the cpSSR play an important role in cp genome rearrangement during evolution (Yap et al. 2015). In this study, we identified 77 cpSSR in the *P. glaucum*

cp genome (Table S4). These abundant SSR loci could provide useful genetic information and sequence resources for further molecular genetic studies in the Poaceae family, including assessment of species genetic diversity and their evolution studies.

The IR/SC boundary regions have been considered to be the most conserved regions in the cp genome (Tsuruta et al. 2017). The contraction and expansion at the borders of IR regions are common evolutionary events in the cp genome, which leads to the major rearrangements and size variations (Raubeson et al. 2007). To find the comprehensive overview between the IR/SC boundary organizations, we compared the *P. glaucum* cp genome with all available Poaceae cp genome. The IR/SC boundaries only showed slight differences, such as contraction and expansion at the borders of IR regions which is relatively common phenomenon in cp genomes (Curci et al. 2015; Daniell et al. 2016; Yang et al. 2013). Moreover, the members of the Poaceae also have lost *accD*, *ycf1*, and *ycf2*, which are absent in the cp genomes of other Panicoideae grasses (Nah et al. 2016; Saski et al. 2007; Tsuruta et al. 2017; Young et al. 2011). The multiple genome alignment indicated that the IR/SC regions were more conserved among the Poaceae cp genomes.

Chloroplast genomes sequencing have been considered as a new tools for phylogenetic analysis with abundant genomic resource in many studies (Cao et al. 2017; Liu et al. 2016). In the present study, we conducted phylogenetic analysis on the 18 Poaceae species and found that the Pearl millet clustered with *S. viridis* and *S. italica* with 100% bootstrap supports. However, the results suggested that additional large-scale genomic analyses using numerous accurately identified Poaceae species are required to clarify the taxonomy and phylogenetic relationships of Millet species at low taxonomic levels.

Acknowledgements

This study was carried out with the support of the “Research Program for Agricultural Science & Technology Development (Project No. PJ008623)” and was supported by the 2017 Postdoctoral Fellowship Program of the National Institute of Agricultural Science, Rural Development Administration, Republic of Korea.

References

- Allen, G.C., Flores-Vergara, M.A., Krasynanski, S., Kumar, S., Thompson, W.F. 2006. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat. Protoc.* **1**:2320–2325.
- Burton, G.W. 1983 Breeding pearl millet. *Plant Breed. Rev.* **1**:162–182.
- Cannarozzi, G.M., Schneider, A. 2012. Codon evolution : mechanisms and models. *Trends Evol. Biol.* **4**:e8.
- Cao, X., Wang J., Wang, H., Liu, S., Chen, L., Tian, X., Qin, H., Wang, L., Na, X., Qiao, Z. 2017. The complete chloroplast genome of *Panicum miliaceum*. *Mitochondrial DNA Part B* **2**:43–45.
- Ceresini, P.C., Silva, C.L.S.P., Missio, R.F., Souza, E.C., Fischer, C.N., Guillherme, I.R., Gregorio, I., Silva, E.H.T., Cicarelli, R.M.B., Silva, M.T.A., Garcia, J.F., Avelar, G.A., Porto Neto, L.R., Marçon, A.R., Bacci Junior, M., Marini, D.C. 2005. Satellyptus: analysis and database of microsatellites from ESTs of *Eucalyptus*. *Genet. Mol. Biol.* **28**:589–600.
- Curci, P.L., De Paola, D., Danzi, D., Vendramin, G.G., Sonnante, G. 2015. Complete chloroplast genome of the multifunctional crop globe artichoke and comparison with other Asteraceae. *PLoS One* **10**:e0120589.

- Daniell, H., Lin, C.S., Yu, M., Chang, W.J. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol.* **17**:134.
- Darling, A.C.E., Mau, B., Blattner, F.R., Perna, N.T. 2004. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **14**:1394–1403.
- Gao, L., Yi, X., Yang, Y.X., Su, Y.J., Wang, T. 2009. Complete chloroplast genome sequence of a tree fern *Alsophila spinulosa*: insights into evolutionary changes in fern chloroplast genomes. *BMC Evol. Biol.* **9**:130.
- Govindaraj, M., Selvi, B., Prabhu, D.A., Rajarathinam, S. 2009. Genetic diversity analysis of pearl millet (*Pennisetum glaucum* [L.] R. Br.) accessions. *Afr J. Biotechnol.* **8**:6046–6052.
- Hu, Z.B., Mbacke, B., Perumal, R. Guèye, M.C., Sy, O., Bouchet, S., Vara Prasad, P.V., Morris, G.P. 2015. Population genomics of pearl millet (*Pennisetum glaucum* (L.) R. Br.): Comparative analysis of global accessions and Senegalese landraces. *BMC Genomics* **16**:ARTN 1048.
- Huang, X., Kurata, N., Wei, X., Wang, Z.X., Wang, A., Zhao, Q., Zhao, Y., Liu, K., Lu, H., Li, W., Guo, Y., Lu, Y., Zhou, C., Fan, D., Weng, Q., Zhu, C., Huang, T., Zhang, L., Wang, Y., Feng, L., Furuumi, H., Kubo, T., Miyabayashi, T., Yuan, X., Xu, Q., Dong, G., Zhan, Q., Li, C., Fujiyama, A., Toyoda, A., Lu, T., Feng, Q., Qian, Q., Li, J., Han, B. 2012. A map of rice genome variation reveals the origin of cultivated rice. *Nature* **490**:497–501.
- Huang, Y.Y., Cho, S.T., Haryono, M., Kuo, C.H. 2017. Complete chloroplast genome sequence of common bermudagrass (*Cynodon dactylon* (L.) Pers.) and comparative analysis within the family Poaceae. *PLoS One* **12**:e0179055.
- ICRISAT, 1996. FAO: The World Sorghum and Millet Economies: Facts, Trends and Outlook vol 68. International Crops Research Institute for the Semi-Arid Tropics, and Rome, Italy: Food and Agricultural Organization of the United Nations, Patancheru, India.
- Kahlau, S., Aspinall, S., Gray, J.C., Bock, R. 2006. Sequence of the tomato chloroplast DNA and evolutionary comparison of solanaceous plastid genomes. *J. Mol. Evol.* **63**:194–207.
- Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., Salzberg, S.L. 2004. Versatile and open software for comparing large genomes. *Genome Biol.* **5**:R12.
- Liu, C.J., Witcombe, J.R., Hash, C.T., Busso, C.S., Pittaway, T.S., Nash, M., Gale, M.D. 1994. Use of molecular marker in sorghum and pearl millet breeding for developing countries. Overseas Development Administration, London U.K.
- Liu, F., Tembrock, L.R., Sun, C., Han, G., Guo, C., Wu, Z. 2016. The complete plastid genome of the wild rice species *Oryza brachyantha* (Poaceae). *Mitochondrial DNA Part B* **1**:218–219.
- Liu, F., Zhao, Y., Luo, D., Hong, D., Li, R. 2017. The complete chloroplast genome sequence of *Oryza eichingeri* (Poaceae). *Mitochondrial DNA Part B* **2**:465–466.
- Lohse, M., Drechsel, O., Kahlau, S., Bock, R. 2013. OrganellarGenomeDRAW – a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* **41**:W575–W581.
- Nah, G., Im, J.H., Lim, S.H., Kim, K.H., Choi, A.Y., Yook, M.J., Kim, S., Kim, C., Kim, D.S. 2016. Complete chloroplast genomes of two Miscanthus species. *Mitochondrial DNA Part A* **27**:4359–4360.
- Olmstead, R.G., Palmer, J.D. 1994. Chloroplast DNA Systematics – a Review of Methods and Data-Analysis. *Am. J. Bot.* **81**:1205–1224.
- Poncet, V., Lamy, F., Enjalbert, J., Joly, H., Sarr, A., Robert, T. 1998. Genetic analysis of the domestication syndrome in pearl millet (*Pennisetum glaucum* L., Poaceae): inheritance of the major characters. *Heredity* **81**:648–658.
- Rajaram, V., Nepolean, T., Senthilvel, S., Varshney, R.K., Vadez, V., Srivastava, R.K., Shah, T.M., Supriya, A., Kumar, S.K., Kumari, B.R., Bhanuprakash, A., Narasu, M.L., Riera-Lizarazu, O., Hash, C.T. 2013. Pearl millet [*Pennisetum glaucum* (L.) R. Br.] consensus linkage map constructed using four RIL mapping populations and newly developed EST-SSRs. *BMC Genomics* **14**:159.
- Ramakrishnan, M., Antony Ceasar, S., Duraipandiyam, V., Al-Dhabi, N.A., Ignacimuthu, S. 2016. Assessment of genetic diversity, population structure and relationships in Indian and non-Indian genotypes of finger millet (*Eleusine coracana* (L.) Gaertn) using genomic SSR markers. *SpringerPlus* **5**:1–11.

- Raubeson, L.A., Peery, R., Chumley, T.W., Dziubek, C., Fourcade, H.M., Boore, J.L., Jansen, R.K. 2007. Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genomics* **8**:174.
- Saarela, J.M., Wysocki, W.P., Barrett, C.F., Soreng, R.J., Davis, J.I., Clark, L.G., Kelchner, S.A., Pires, J.C., Edger, P.P., Mayfield, D.R., Duvall, M.R. 2015. Plastid phylogenomics of the cool-season grass subfamily: clarification of relationships among early-diverging tribes. *AoB PLANTS* **7**:plv046.
- Saski, C., Lee, S.B., Fjellheim, S., Guda, C., Jansen, R.K., Luo, H., Tomkins, J., Rognli, O.A., Daniell, H., Clarke, J.L. 2007. Complete chloroplast genome sequences of *Hordeum vulgare*, *Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with other grass genomes. *Theor. Appl. Genet.* **115**:591.
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A., Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**:2725–2729.
- Thiel, T., Michalek, W., Varshney, R.K., Graner, A. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **106**:411–422.
- Tsuruta, S., Ebina, M., Kobayashi, M., Takahashi, W. 2017. Complete chloroplast genomes of *Eriarhynchus arundinaceus* and *Miscanthus sinensis*: Comparative genomics and evolution of the saccharum complex. *PLoS One* **12**:e0169992.
- Wu, P., Zhou, C., Cheng, S., Wu, Z., Lu, W., Han, J., Chen, Y., Chen, Y., Ni, P., Wang, Y., Xu, X., Huang, Y., Song, C., Wang, Z., Shi, N., Zhang, X., Fang, X., Yang, Q., Jiang, H., Chen, Y., Li, M., Wang, Y., Chen, F., Wang, J., Wu, G. 2015. Integrated genome sequence and linkage map of physic nut (*Jatropha curcas* L.), a biodiesel plant. *Plant J.* **81**:810–821.
- Wyman, S.K., Jansen, R.K., Boore, J.L. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* **20**:3252–3255.
- Yang, J.B., Tang, M., Li, H.T., Zhang, Z.R., Li, D.Z. 2013. Complete chloroplast genome of the genus *Cymbidium*: lights into the species identification, phylogenetic implications and population genetic analyses. *BMC Evol. Biol.* **13**:84.
- Yap, J.Y., Rohner, T., Greenfield, A., Van Der Merwe, M., McPherson, H., Glenn, W., Kornfeld, G., Marendy, E., Pan, A.Y., Wilton, A., Wilkins, M.R., Rossetto, M., Delaney, S.K. 2015. Complete chloroplast genome of the wollemi pine (*Wollemia nobilis*): Structure and evolution. *PLoS One* **10**:e0128126.
- Young, H.A., Lanzatella, C.L., Sarath, G., Tobias, C.M. 2011. Chloroplast genome variation in upland and lowland switchgrass. *PLoS One* **6**:e23980.
- Zhang, Y., Nie, X., Jia, X., Zhao, C., Biradar, S.S., Wang, L., Du, X., Weining, S. 2012. Analysis of codon usage patterns of the chloroplast genomes in the Poaceae family. *Aust. J. Bot.* **60**:461–470.
- Zhou, M., Long, W., Li, X. 2008. Patterns of synonymous codon usage bias in chloroplast genomes of seed plants. *For. Stud. China* **10**:235.

Electronic Supplementary Material (ESM)

Electronic Supplementary Material (ESM) associated with this article can be found at the website of CRC at <https://akademai.com/loi/0806>

Electronic Supplementary *Table S1*. Gene composition in the *P. glaucum* chloroplast genome

Electronic Supplementary *Table S2*. The relative synonymous codon usage in the *P. glaucum* chloroplast genome

Electronic Supplementary *Table S3*. The calculated parameter values for each coding genes in Pearl millet

Electronic Supplementary *Table S4*. The candidate cpSSR in the Pearl millet chloroplast genome

Electronic Supplementary *Figure S1*. Locally collinear blocks (LCBs) analysis of nine Poaceae chloroplast genomes. The *A. sativa* cp genome is shown at top as the reference genome. LCBs are indicated by the thin colored lines. Histograms within each block represent the degree of sequence similarity. LCBs lying below the black line have an inverse orientation. The small boxes below each chloroplast genome indicate genes; upper and lower boxes are transcribed counterclockwise and clockwise, respectively. Red boxes indicate ribosomal RNA genes

Electronic Supplementary *Figure S2*. Maximum likelihood (ML) phylogeny of the 18 Poaceae species using complete whole chloroplast genome sequences. Numbers above each node indicate the ML bootstrap support values. The branches without numbers indicate 100% bootstrap supports