Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

OPEN ACCESS   Check for updates

The plastid genome of *Pentadiplandra brazzeana* Baillon (Pentadiplandraceae)

Xu Zheng[a,b] and Ting-Shuang Yi[b]

[a]College of Life Sciences, Inner Mongolia Agricultural University, Hohhot, China; [b]Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China

ABSTRACT

Pentadiplandra brazzeana is the only species of Pentadiplandraceae. This species is shrubs or tuberous liana with a natural distribution in tropical West Africa. This study firstly determined its complete plastome. The plastome is totally 156,625 bp in length, which contains a pair of 26,751-bp-long inverted repeat regions (IRs), a large single copy region of 85,318 bp, and a small single copy region of 17,805 bp. A total of 112 unique genes were identified in this plastome, of which 78 are protein-coding genes, 30 are tRNA genes, and 4 are rRNA genes. Phylogenetic analysis based on 82 genes fully resolved relationships among sampled families of Brassicales, which are consistent with previous studies.

Pentadiplandra brazzeana is the only species of the family Pentadiplandraceae (Hall et al. 2004). It is an evergreen shrub or liana that is found in tropical West Africa, among northern Angola, eastern Nigeria and western Democratic Republic of Congo. It produces large red berries, sometimes mottled with gray. Its fruits contain two kind of high-potency thermostable sweet protein, called brazzein (Ming and Hellekant 1994) and pentadin (Vanderwel et al. 1989), which make the berry substantially sweeter than saccharose. Due to its sweetness, heat stability, high water solubility and minimum molecular weight, brazzein is the most superior protein sweetener presently known (Ming and Hellekant 1994). In this study, we determined the complete plastome of *P. brazzeana* based on the whole-genome Illumina sequencing dataset. We have deposited the annotated plastome of *P. brazzeana* in the GenBank with the accession number MN306574. Being the first complete plastome data for *P. brazzeana* as well as the family Pentadiplandraceae, the plastome sequence could be used for conservation and genetics studies of this endemic species, and the phylogenetic analysis of interfamilial relationships of the whole Brassicales.

The material of *P. brazzeana* used for this project has been collected from Kisangani, Democratic Republic of the Congo(0°31′0″N 25°12′0″E), and its specimen (Druk. 281/174) is deposited in the Royal Botanic Gardens, Kew. Total genomic DNA was obtained from the Royal Botanic Gardens, Kew. Paired-end sequencing was performed on Illumina HiSeq X TEN at Plant Germplasm and Genomics Center (Kunming Institute of Botany). The paired-end reads were filtered and assembled into a complete plastome using GetOrganelle v. v1.6.2a (Camacho et al. 2009; Bankevich et al. 2012; Langmead and Salzberg 2012; Jin et al. 2018), with final assembly graph checked using Bandage (Wick et al. 2015). The plastome was automatically annotated using PGA (Qu et al. 2019) and GeSeq (Tillich et al. 2017), then manually adjusted in Geneious 9.0.5 (Kearse et al. 2012). The plastome of *P. brazzeana* is 156,625 bp in length, containing a large single copy (LSC) region of 85,318 bp and a small single copy (SSC) region of 17,805 bp separated by a pair of 26,751-bp-long IR regions. The overall GC content was 36.3% and the plastome contains 112 unique genes (78 protein-coding genes, 30 tRNA genes, and 4 rRNA genes). Seven protein-coding, seven tRNA, and four rRNA genes are duplicated in the IR regions, the *ndhK* gene was pseudogene.

In order to verify the phylogenetic relationships of *P. brazzeana* in Brassicales, we constructed a maximum-likelihood (ML) tree of six important families of Brassicales that have plastome sequences, three representative species for Brassicaceae, and each one species for other five families. The *Vitis rotundifolia* (Vitaceae) and *Bixa orellana* (Bixaceae) were used as outgroup (Figure 1). Phylogenetic analysis was performed on a data set including 82 gene sequences using RAxML version 8.2.11 with 1000 bootstrap replicates (Stamatakis 2014). Relationships among sampled families of Brassicales were fully resolved with strong supports at most nodes, which are consistent with previous studies (Edger et al. 2018).
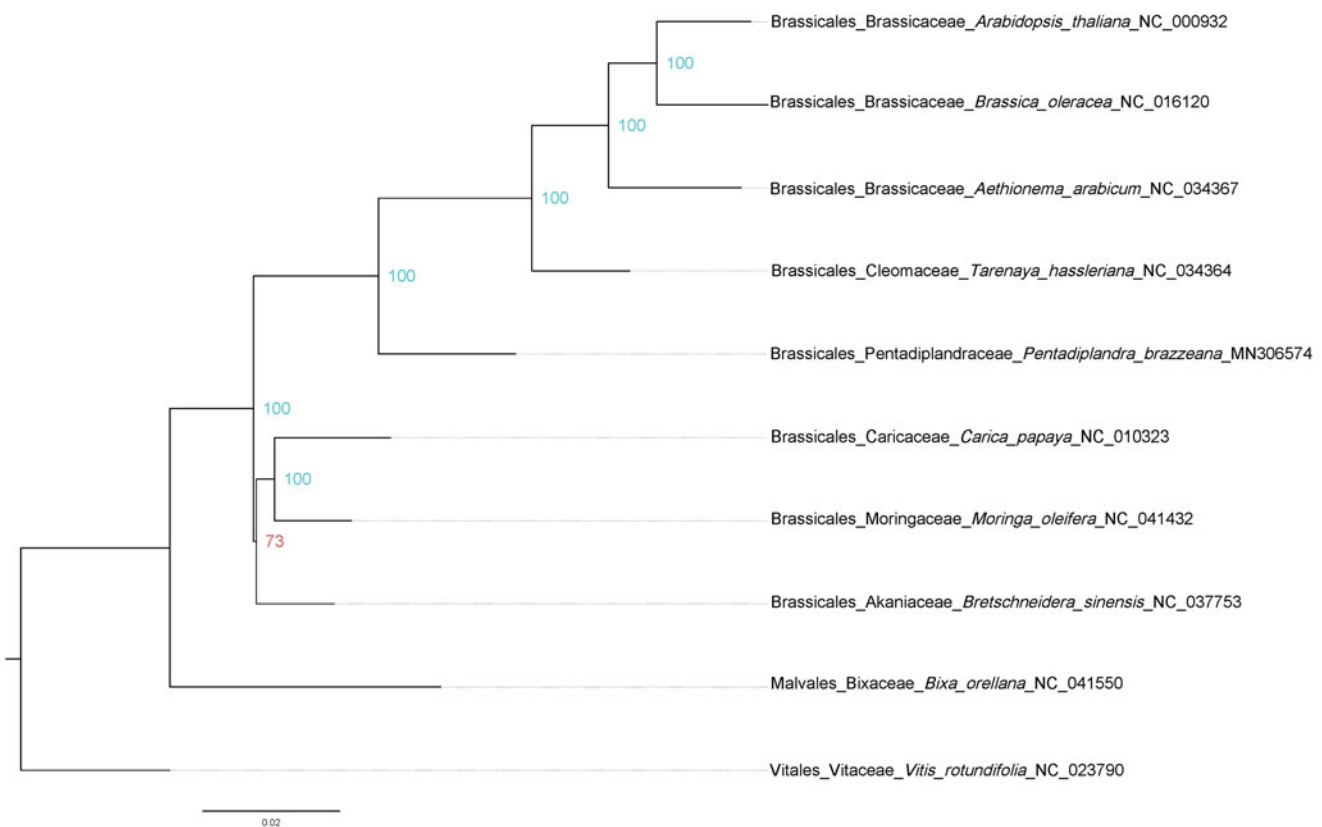
Figure 1. Maximum-likelihood (ML) tree inferred from the 82-gene (78 protein-coding and 4 rRNA genes) matrix. Numbers at the right of nodes are bootstrap support values. GenBank accession numbers: *Arabidopsis thaliana* (NC_000932), *Brassica oleracea* (NC_016120), *Aethionema arabicum* (NC_034367), *Tarenaya hassleriana* (NC_034364), *Pentadiplandra brazzeana* (MN306574), *Carica papaya* (NC_010323), *Moringa oleifera* (NC_041432), *Bretschneidera sinensi* (NC_037753), *Bixa orellana* (NC_041550), and *Vitis rotundifolia* (NC_023790).

References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5):455–477.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST plus: architecture and applications. BMC Bioinformatics. 10(1):421.

Edger PP, Hall JC, Harkess A, Tang M, Coombs J, Mohammadin S, Schranz ME, Xiong ZY, Leebens-Mack J, Meyers BC, et al. 2018. Brassicales phylogeny inferred from 72 plastid genes: a reanalysis of the phylogenetic localization of two paleopolyploid events and origin of novel chemical defenses. Am J Bot. 105(3):463–469.

Hall JC, Iltis HH, Sytsma KJ. 2004. Molecular phylogenetics of core brassicales, placement of orphan genera Emblingia, Forchhammeria, Tirania, and character evolution. Syst Bot. 29(3):654–669.

Jin JJ, Yu WB, Yang JB, Song Y, Yi TS, Li DZ. 2018. GetOrganelle: a simple and fast pipeline for de novo assembly of a complete circular chloroplast genome using genome skimming data. bioRxiv. 256479.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28(12): 1647–1649.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods. 9(4):357–359.

Ming D, Hellekant G. 1994. Brazzein, a new high-potency thermostable sweet protein from *Pentadiplandra brazzeana-B*. FEBS Lett. 355(1): 106–108.

Qu XJ, Moore MJ, Li DZ, Yi TS. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. Plant Methods. 15:50.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9): 1312–1313.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq – versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45:6–11.

Vanderwel H, Larson G, Hladik A, Hladik CM, Hellekant G, Glaser D. 1989. Isolation and characterization of Pentadin, the sweet principle of *Pentadiplandra brazzeana* Baillon. Chem Senses. 14:75–79.

Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics. 31(20): 3350–3352.