Conservation Genet Resour (2018) 10:405–407
DOI 10.1007/s12686-017-0836-8

TECHNICAL NOTE

The chloroplast genome of a rare and an endangered species *Salweenia bouffordiana* (Leguminosae) in China

Hafiz Muhammad Wariss[1,2,3] · Ting-Shuang Yi[2] · Hong Wang[1] · Rong Zhang[2,3]

Received: 17 July 2017 / Accepted: 9 August 2017 / Published online: 18 August 2017
© Springer Science+Business Media B.V. 2017

Abstract Salweenia bouffordiana is an endangered species of Qinghai-Tibetan Plateau (QTP) in China. In this study, its complete plastome was assembled using next-generation sequencing. The complete genome is 153, 730 bp in length, comprises a pair of inverted repeat (IR; 26, 227 bp) regions separated by the large single-copy (LSC; 83, 509 bp) and small single-copy (SSC; 17, 767 bp) regions. It encodes 111 genes, including 77 protein-coding genes (PCGs), 30 tRNAs and 4 rRNAs. The nucleotide composition is asymmetric (31.6% A, 18.1% C, 18.7% G and 31.6% T) with an overall A+T content of 73.2%. Three types of 111 simple sequence repeats (SSRs) were detected in the plastome, including 107 mono- and 4 di-nucleotide repeats. The newly sequenced complete plastome and identified SSRs will help in understanding the plastome evolution and genetic conservation of *S. bouffordiana*.

Keywords *Salweenia bouffordiana* · Endangered species · Plastome · SSR · Leguminosae

Salweenia Baker f. is a Tertiary relict genus within the Legumes (Li and Ni 1982; Wang 1989; Wu et al. 2003). It was first described by Baker (1935), confined to dry shrub lands and croaky slopes of the Lancang Jiang, Nu Jiang, and Yalong Jiang valleys of QTP in China (Li and Ni 1982; Ding 1994). *Salweenia bouffordiana* H. Sun, Z.M. Li & J.P. Yue is an evergreen shrub with beautiful yellow flowers, which occurs within a very narrow range of a fragile habitat in Yalong Jiang, a tributary of the Jinsha Jiang (Yue et al. 2011). Currently, due to the anthropogenic activities, its population is on the verge of rapid decline. The species is critically endangered (CR) (Yue et al. 2011) following the IUCN guidelines (IUCN 2012) and belongs to the rare and endangered plants listed in the category of key protected wild plants in China (http://www.zhb.gov.cn/gkml/hbb/bgg/201309/t20130912_260061.htm). Considering the importance of the preservation of the precious germplasm, it is necessary to explore available genetic resource for this species. However, chloroplast genome supplies useful molecular markers to reveal genetic diversification, the basis for designing efficient conservation strategies for endangered species (Mukherjee et al. 2016; Twyford and Ness 2017).

DNA material of *S. bouffordiana* (voucher specimens: Yi380) was obtained from the Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences. Total genomic DNA was sent to Beijing Genomics Institute, Shenzhen, China for constructing a paired-end (PE) library and sequencing using the Illumina HiSeq2000. The plastome was filtered and assembled using GetOrganelle.py pipeline (https://github.com/Kinggerm/GetOrganelle) with reference (*Arachis hypogaea* L., GenBank: NC026676). Contigs were connected into complete plastome using Bandage Linux dynamic v.8.0 (Wick et al. 2015). PE reads were mapped to the plastome using Bowtie2 (Langmead and Salzberg 2012). Then, annotation of plastome was performed

Electronic supplementary material The online version of this article (doi:10.1007/s12686-017-0836-8) contains supplementary material, which is available to authorized users.

✉ Rong Zhang
zhangronga@mail.kib.ac.cn

[1] Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, China

[2] Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, China

[3] University of Chinese Academy of Sciences, Beijing, China

on GeSeq (Tillich et al. 2017), coupled with adjustment in Geneious v.9.1.4 (Kearse et al. 2012). Finally, the circular genome map was generated with OGDRAW (Lohse et al. 2013). The annotated plastome of *S. bouffordiana* has been submitted to GenBank (MF449303). SSRs were detected using MISA perl script (http://pgrc.ipk-gatersleben.de/misa/misa.html), with the threshold: ten repeat units for mononucleotide SSRs, six repeat units for dinucleotide, five repeat units for tri-, tetra-, penta- and hexa-nucleotide SSRs.

The complete plastome of S. bouffordiana is circular with 153, 730 bp in length, and contains a pair of inverted repeat (IR) regions of 26, 227 bp each, separated by a LSC region of 83, 509 bp and a SSC region of 17, 767 bp (Fig. 1). It encodes 111 genes, including 77 PCGs, 30 tRNAs and four rRNAs. Among these, nine genes (*atpF, ndhA, ndhB, petB,*

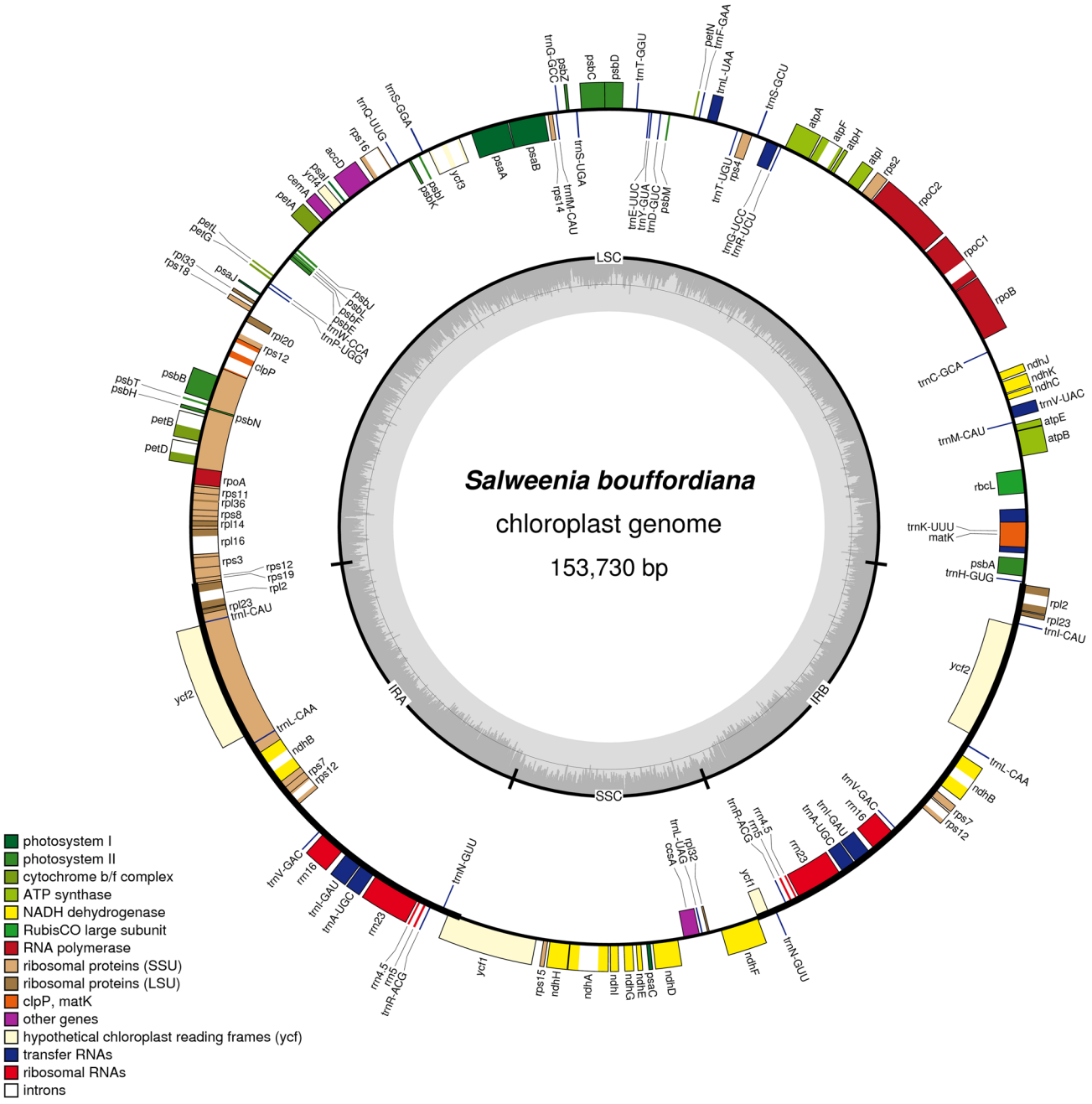

Fig. 1 Gene map of the complete plastome of Salweenia bouffordiana. Genes are indicated by *boxes* on the inside (clockwise transcription) and outside (counterclockwise transcription) of the outermost circle. Genes belonging to different functional groups are *color* coded. *Dashed area* in the inner circle indicates the GC content of the plastome. (Color figure online)

Table 1 Simple sequence repeats (SSRs) detected in the chloroplast genome of *Salweenia bouffordiana*

Repeats	Number	Location			Region		
		LSC	IR	SSC	CDS	Intron	IGS
A/T	96	66	8	22	19	15	46
C/G	11	4	6	1	0	4	7
AT	4	4	0	0	0	0	4
Sum	111	74	14	23	19	19	57

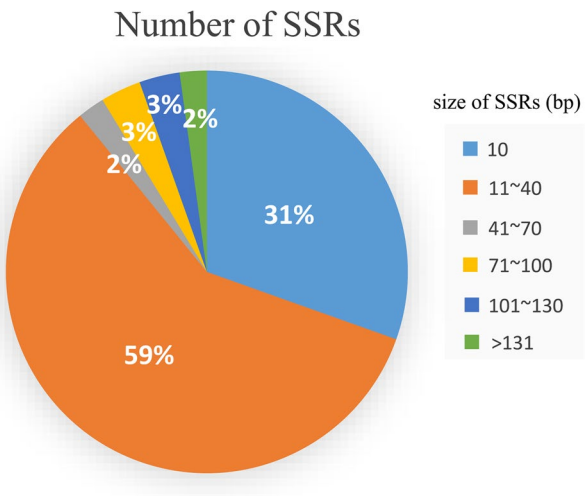

Fig. 2 Distribution percentage of predicted SSRs in different lengths

petD, rpl2, rpl16, rpoC1 and *rps16*) harbor a single intron while three genes (*clpP, rps12* and *ycf3*) possess two introns. Seven PCGs (*ndhB, rpl2, rpl23, rps7, rps12, ycf1* and *ycf2*), 7 tRNAs genes (*trnA-UGC, trnI-CAU, trnI-GAU, trnL-CAA, trnN-GUU, trnR-ACG* and *trnV-GAC*) and 4 rRNAs genes (*rrn4.5, rrn5, rrn16* and *rrn23*) are duplicated in both IR regions (Supplementary file 1). The overall GC content was 36.8%, lower than that of the IR regions (42.6%), but higher than those of the LSC (34.5%) and SSC (30.5%) regions.

Within the complete plastome of *S. bouffordiana*, there are 111 SSRs with a l ength of at least 10 bp, including 107 mono- and 4 di-nucleotide repeats (Table 1). However, no tri-, tetra-, penta- or hexanucleotide repeats were found. Majority (96/107) of the mononucleotides consist of A or T and all of the dinucleotides are AT repeats. This is consistent with the AT-richness of the chloroplast genome. Nineteen SSRs are located in 6 PCRs (*atpB, ndhK, rpoB, rpoC2, ycf1* and *ndhF*) while 19 are located in introns (introns of *rpoC1, atpF, ycf3, rps16, clpP, petD, rpl16, ndhA* and *trnK-UUU*) of the plastome (Supplementary file 2). The size of SSRs were classified into six groups (Fig. 2). The shortest SSRs (10 bp) constituted 31%, whereas SSRs with lengths of 11–40 and 41–70 bp accounted for 59 and 2%, respectively.

Acknowledgements We appreciate funding from the Talent Project of Yunnan Province (No. 2011CI042) and are grateful to the Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences for providing DNA materials.

Funding This study was funded by grant from the Talent Project of Yunnan Province (No. 2011CI042).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Baker EG (1935) New genus of Leguminosae from Tibet. J Bot 73:134–136

Ding CS (1994) Robinieae. In: Wei Z (ed) Flora Reipublicae Popularis Sinicae, vol 40. Science Press, Beijing, pp 226–228

IUCN (2012). IUCN Red List Categories and Criteria, version 3.2. Prepared by the IUCN Species Survival Commission. IUCN, Gland

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierern T, Ashton B, Meintjes P, Dremmond A (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649

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

Li PC, Ni ZC (1982) The formation and differentiation of the Leguminosae flora in Xizang (Tibet). Acta Phytotaxon Sin 20:142–156

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

Mukherjee A, Williams D, Gitzendanner MA, Overholt WA, Cuda JP (2016) Microsatellite and chloroplast DNA diversity of the invasive aquatic weed Hygrophila polysperma in native and invasive ranges. Aquat Bot 129:55–61. doi:10.1016/j.aquabot.2015.12.004

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 doi:10.1093/nar/gkx391

Twyford AD, Ness RW (2017) Strategies for complete plastid genome sequencing. Mol Ecol Resour. doi:10.1111/1755-0998.12626

Wang HS (1989) A study on the origin of spermatophytic genera endemic to China. Acta Bot Yunnan 11:1–16

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

Wu ZY, Lu AM, Tang YC, Chen ZD, Li DZ (2003) The families and genera of angiosperms in China: a comprehensive analysis. Science Press, Beijing

Yue XK, Yue JP, Yang LE, Li ZM, Sun H (2011) Systematics of the genus *Salweenia* (Leguminosae) from Southwest China with discovery of a second species. Taxon 60:1366–1374