

## MITOGENOME ANNOUNCEMENT

**The complete chloroplast genome sequence of *Perilla frutescens* (L.)**Qi Shen<sup>1,2</sup>, Jun Yang<sup>1</sup>, Chaolong Lu<sup>1,3</sup>, Bo Wang<sup>1,4</sup>, and Chi Song<sup>1</sup>

<sup>1</sup>Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, P.R. China, <sup>2</sup>Guizhou Rapeseed Institute, Guizhou Province of Academy of Agricultural Sciences, Guiyang, P.R. China, <sup>3</sup>Maize Research Institute, Sichuan Agricultural University, Wenjiang, P.R. China, and <sup>4</sup>Institute of Medicinal Plant Development, Chinese Academy of Medical Science & Peking Union Medical College, Beijing, P.R. China

**Abstract**

*Perilla frutescens* (L.) is a traditionally medical herb of East Asia. The complete chloroplast genome of *P. frutescens* (L.) Britton var. *frutescens* was assembled in this study. Total chloroplast genome size of *Perilla* was 153,666 bp in length, containing a pair of inverted repeats (IRs) of 25,677 bp, separated by large single copy (LSC) and small single copy (SSC) of 84,288 bp and 17,925 bp, respectively. Overall GC contents of the genome were 37.6%. The chloroplast genome harbored 127 annotated genes, including 89 protein-coding genes, 29 tRNA genes and 8 rRNA genes. Eleven genes contained one or two introns.

**Keywords**

Chloroplast, genome sequence, medicinal plant *Perilla frutescens* (L.)

**History**

Received 14 January 2015

Accepted 24 January 2015

Published online 25 February 2015

*Perilla frutescens* (L.) is one of valuable economical crop and traditionally medical herb of Lamiaceae family (Park et al., 2008). *Perilla* plants are widespread cultivated in East Asian countries, including Korea, China, and Japan in the past years (Park et al., 2008). Recently, diverse utilization of *Perilla* plants had been developed. Its seeds were produced foods and nutritional edible oil, fresh leaves were used as vegetable and flavoring, leaves also were used as traditional pot-herb for medicine or food coloring (Pandey & Bhatt, 2008; You et al., 2014). Chloroplast is a semiautonomous organelle, which possesses its own genome that codes some of genes involving in photosynthesis and house-keeping function. Chloroplast genomes are non-meiotic and mostly uniparental inheritance. These characteristics render them as important informational standards in plant phylogenetic studies (Huang et al., 2014). In the article, complete chloroplast genome sequence of *Perilla frutescens* (L.) Britton var. *frutescens*, which is an endemic genus in China, were sequenced and characterized.

Total genome DNA of *Perilla* was extracted to construct 180bp small-size fractionated libraries according to NEBNext Ultra DNA library Prep Kit for Illumina. Whole genome sequences were executed using Illumina HiSeq 1500 Genome Analyzer (rapid run model configuration for pair-end 101 reads) at institute of chinese materia medica in China. Raw reads were filtered by standard method (Patel & Jain, 2012). Then, 7,823,427 filtered reads were assembled into the complete chloroplast

genome using SOAP denovo software (Luo et al., 2012) with *Salvia miltiorrhiza* chloroplast genome sequence as reference genome (Qian et al., 2013). The chloroplast genome annotation was according to methods described by Liu et al. (2012). The most homologous of protein, cDNA and rRNA from GenBank were allocated on *Perilla* chloroplast genome with the combination of Blastx, Blastn, protein2genome and est2genome programs (Liu et al., 2012). And the inverted repeats were identified using the vmatch software tool (Abouelhoda et al., 2004), tRNAs were identified using tRNAscan (Lowe & Eddy, 1997). The complete chloroplast genome sequence together with gene annotations were submitted to the GenBank under the accession number of KP404631.

Complete chloroplast genome of *Perilla* has a total length of 153,666 bp. Similarly to typical chloroplast genome structure in higher plants, it has a large single copy (LSC) region of 80,468 bp and a small single copy (SSC) region of 13,748 bp, is separated by a pair of inverted repeats (IRa and IRb) of 25,677 bp. The GC content of the whole chloroplast genome was 37.6%, with the IRs has a higher GC content (43.1%) than LSC (35.6%) and SSC (31.2%) region due to the presence of GC-rich rRNA genes. The gene coding regions account for 57.9% of the genome. It includes 127 unique genes, composed of 89 protein-coding genes and 8 rRNA genes. Among these genes, 8 genes contains one intron, and 3 genes have two introns. The chloroplast genome contains 29 different tRNA genes, and 10 are present in the IRs (Figure 1).

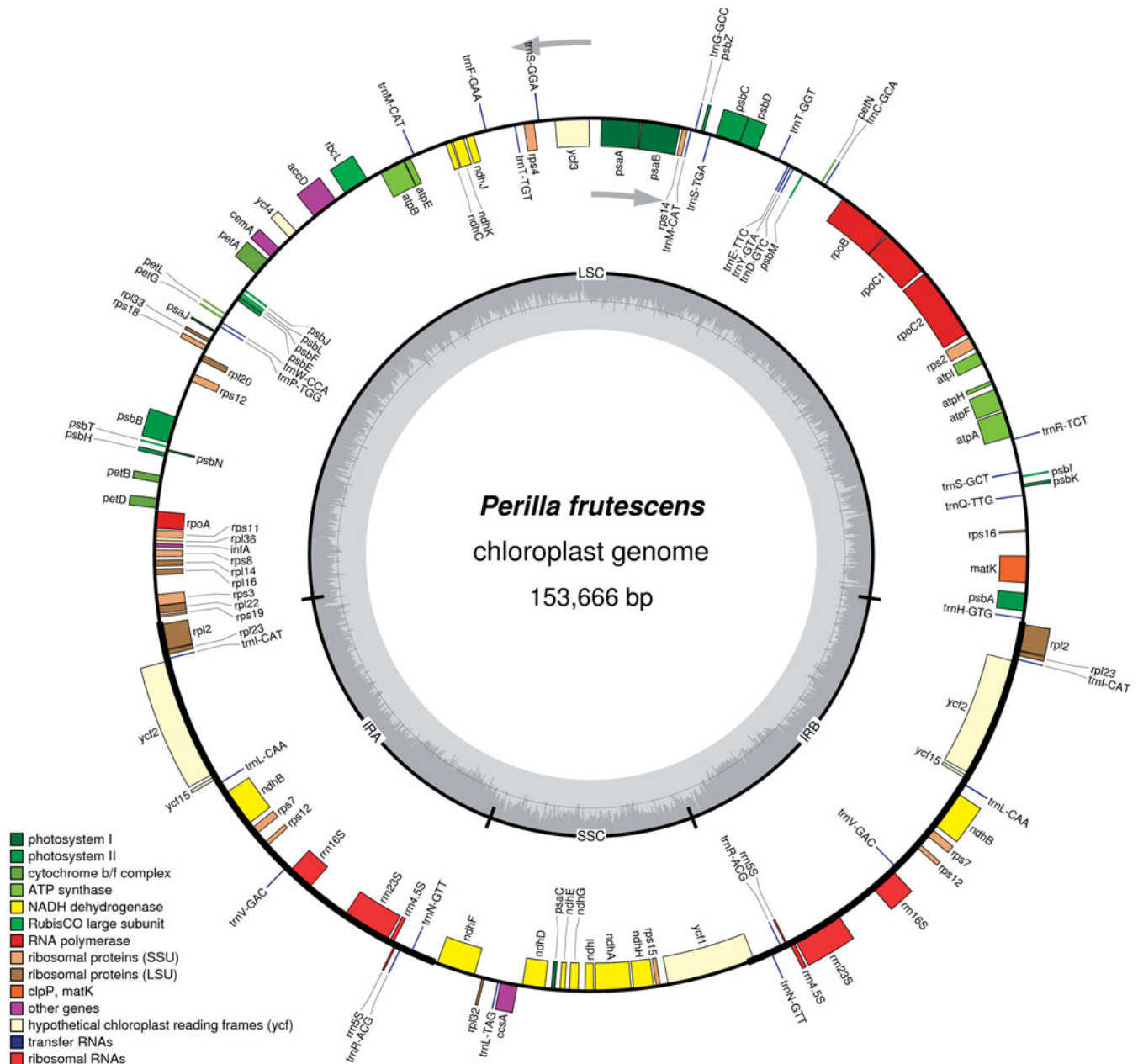


Figure 1. Gene map of *P. frutescens* chloroplast genome. Genes drawn inside the circle are transcribed clockwise, while those drawn outside the circle are counterclockwise. Genes belonging to different functional groups are color-coded. The darker gray in the inner circle corresponds to the GC content.

## Declaration of interest

This work was supported by China Postdoctoral Science Foundation (2014M550135) and National Natural Science Foundation of China (31360067).

## References

- Abouelhoda MI, Kurtz S, Ohlebusch E. (2004). Replacing suffix trees with enhanced suffix arrays. *J Dis Algor* 2:53–86.
- Huang H, Shi C, Liu Y, Mao SY, Gao LZ. (2014). Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: Genome structure and phylogenetic relationships. *BMC Evol Biol* 14:151.
- Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X. (2012). CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genomics* 13:715.
- Lowe TM, Eddy SR. (1997). tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:0955–64.
- Luo RB, Liu BH, Xie YL, Li ZY, Huang WH, Yuan JY, He GH, et al. (2012). SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *Giga Sci* 1:18.
- Pandey A, Bhatt K. (2008). Diversity distribution and collection of genetic resources of cultivated and weedy type in *Perilla frutescens* (L.) Britton var. *frutescens* and their uses in Indian Himalaya. *Genet Resour Crop Ev* 55:883–92.
- Park YJ, Dixit A, Ma KH, Lee JK, Lee MH, Chung CS, Nitta M, et al. (2008). Evaluation of genetic diversity and relationships within an on-farm collection of *Perilla frutescens* (L.) Britt. using microsatellite markers. *Genet Resour Crop Ev* 55:523–35.
- Patel RK, Jain M. (2012). NGS QC Toolkit: A toolkit for quality control of next generation sequencing data. *PLoS one* 7:e30619.
- Qian J, Song J, Gao H, Zhu Y, Xu J, Pang X, Yao H, et al. (2013). The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*. *PLoS one* 8:e57607.
- You CX, Yang K, Wu Y, Zhang WJ, Wang Y, Geng ZF, Chen HP, et al. (2014). Chemical composition and insecticidal activities of the essential oil of *Perilla frutescens* (L.) Britt. aerial parts against two stored product insects. *Eur Food Res Technol* 239:481–90.