Korean Journal of Plant Taxonomy
52(2): 123–126, 2022

SHORT COMMUNICATION
https://doi.org/10.11110/kjpt.2022.52.2.123

The complete chloroplast genome sequence of *Hibiscus sabdariffa* (Malvaceae)

Soon-Ho KWON, Yunmi PARK, You Lim JANG and Hae-Yun KWON*

Department of Forest Bio-Resources, National Institute of Forest Science, Suwon 16631, Korea

Corresponding author
Hae-Yun KWON
E-mail: kwonhy05@korea.kr

OPEN ACCESS
© 2022 the Korean Society of Plant Taxonomists.
This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT: *Hibiscus sabdariffa* L., (roselle) in the Malvaceae family is an erect subshrub known to be native to India and Malaysia. It is widely used as a food or tea material around the world, and its therapeutic effects have been widely studied. In this study, the sequencing of the complete chloroplast genome of *H. sabdariffa* was carried out. The result indicates a genome size of 162,428 bp, which is composed of a large single copy of 90,327 bp, two inverted repeats of 26,242 bp each, and a small single copy of 19,617 bp. Overall, a total of 131 genes were predicted, including 86 coding sequences, 37 tRNAs, and 8 rRNAs. According to a phylogenic analysis, it was clearly distinguished from outgroups such as other species of the genus *Hibiscus* used in the analysis.

Keywords: chloroplast genome, edible plant, herbal medicine, *Hibiscus sabdariffa,* phylogenetic analysis, roselle

RECEIVED 19 April 2022; **REVISED** 24 May 2022; **ACCEPTED** 16 June 2022

INTRODUCTION

Malvaceae consists of 244 genera and 4,225 species (Christenhusz and Byng, 2016). In the genus *Hibiscus*, more than 250 species are distributed from the temperate to tropical climate regions (Mandaji et al., 2022), and these species include herbs, shrubs, or trees (Abdullah et al., 2020). Plants of the genus *Hibiscus* have been studied and used not only for ornamental purposes, but also as functional foods and medicines (Abdelhafez et al., 2020).

Hibiscus sabdariffa L., commonly known as roselle, is an herbaceous subshrub and can reach up to 2.5 m in height, and generally possesses red stems and calyces (Morton, 1987; Da-Costa-Rocha et al., 2014; Bule et al., 2020). Roselle is known to be native to India and Malaysia, but it has been cultivated and spread in countries with tropical and subtropical climates regions (Izquierdo-Vega et al., 2020). Roselle has a variety of uses: the seeds are used as a source of dietary fibres, antioxidants, and edible oil. The dried calyx is used in beverages and teas known as carcade that are effective against chronic non-communicable diseases (Dhar et al., 2015; Montalvo-González et al., 2022). Roselle has been mainly studied on its utility in various industries, but few on genetic characteristics and structure (Sánchez-Mendoza et al., 2008; Montalvo-González et al., 2022).

Chloroplast is an organelle that performs primarily photosynthesis, and its genome sequence is well conserved, so it is a major material utilized in the studies of species classification, differentiation and evolutionary process (Liu et al., 2019; Cheng et al., 2020; Kim et al., 2021). In this study, the chloroplast complete genome sequence assembly of roselle was peformed for the first time to be used as an important material for further studies on the evolution and biodiversity of the species and among the species of *Hibiscus.*

MATERIALS AND METHODS

Fresh leaves of *H. sabdariffa* were sampled in National Institute of Forest Science, Suwon, Korea. Total DNA was extracted by GeneAll Genomic DNA Purification Kit (GeneAll Biotechnology, Seoul, Korea). DNA library was constructed using TruSeq Nano DNA Kit with a protocol according to the Sample Preparation Guide provided by the manufacturer (Macrogen Inc., Seoul, Korea). Genome sequencing was performed on the Illumina NovaSeq 6000 platform with 151 bp read size and paired-end type (Macrogen Inc.).

The chloroplast complete genome was assembled using

NOVOPlasty v.4.3.1, an organelle assembler (Dierckxsens et al., 2017). To increase assembly reliability, repeated assemble work was performed with *K*-mer 21, 23, 25, 27, 29, 31, 33 based on two reference genome sequences, *H. syriacus*, NC_026909 (Kwon et al., 2016) and *H. sinosyriacus*, MZ_367751. Comparative sequence verification and error correction were carried out by manually. Gene annotation was performed using BLATN, BLATX, and Chloe v. 0.1.0 in GeSeq, an annotator of organelle genomes (Tillich et al., 2017). A circular map of the chloroplast complete genome was drawn by OGDRAW v. 1.3.1 (Greiner et al., 2019). Microsatellite analysis was carried out by MIcroSAtellite (MISA) v. 2.1 with a default setting (Beier et al., 2017).

For phylogenic analysis, total nine chloroplast genome sequences of Malvaceae were used, including *H. sabdariffa* and two outgroups, *Gossypium raimondii* and *G. trilobum*. Sequences were aligned using the Clustal Omega v. 1.2.4, alignment program (Sievers et al., 2011). The phylogenetic tree was reconstructed using the Maximum Likelihood method with the JTT matrix-based model and 1,000 bootstrap replicates in MEGA v. 11.0.11 (Tamura et al., 2021).

RESULTS AND DISCUSSION

A total of 159,895,334 reads were produced in assembly work. Of these, 1,768,742 reads, about 9% of total reads mapped to the reference genomes, *H. syriacus* and *H. sinosyriacus* and the average organelle coverage of total reads compared to overlapped sequences was 9162 x . Finally, the chloroplast complete genome of *H. sabdariffa* was assembled. Its total genome size was 162,428 bp. The chloroplast genome of the species was registered in NCBI's GenBank with accession

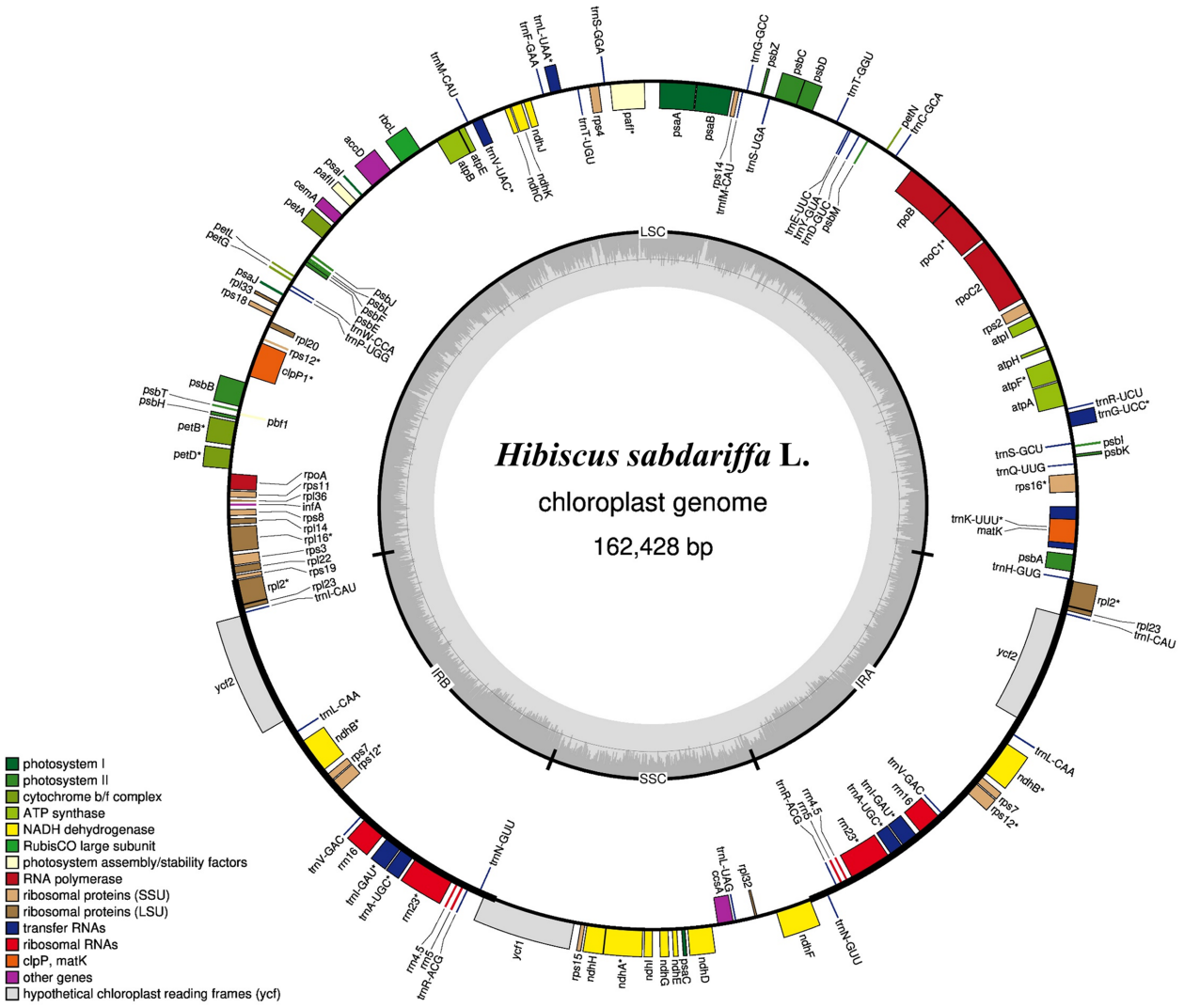

Fig. 1. A circular map and annotations of chloroplast complete genome of *Hibiscus sabdariffa* drawn by OGDRAW v. 1.3.1 and annotated by BLATN, BLATX, and Chloe v. 0.1.0 in GeSeq organelle annotation web tool.

Table 1. Distribution of simple sequence repeats in *Hibiscus sabdariffa*.

Structure	LSC			IRa	SSC	IRb
Repeat type	A/T	AT/TA	AAT	A/T	A/T	A/T
No.	57	3	1	3	9	2

LSC, large single copy; SSC, small single copy. IR, inverted repeat.

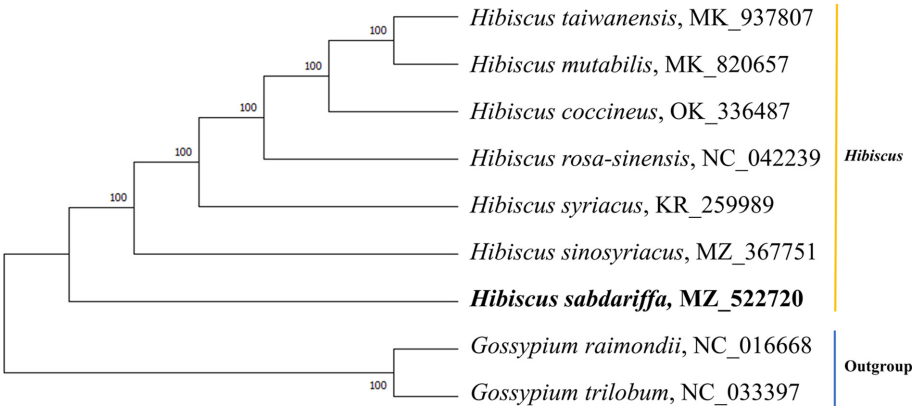

Fig. 2. Phylogenetic relationship of *Hibiscus sabdariffa.* Phylogenetic analysis was carried out via the maximum likelihood method using the MEGA v.11.0.11 program. Bootstrap values, derived from 1,000 pseudoreplicates were indicated near the nodes.

number, MZ_522720. The associated BioProject, BioSample, and SRA numbers are PRJNA_789603, SAMN_24146078, and SRR_17253144, respectively. The genome is composed of four regions: large single-copy (LSC), two inverted repeats (IRs), and small single-copy (SSC). LSC contains 90,327 bp, while IRa and IRb contain 26,242 bp each, and SSC has 19,617 bp (Fig. 1). In total, 131 genes comprising 86 coding sequences, 37 tRNAs, and 8 rRNAs are predicted. There were 75 simple sequence repeats (SSRs): 71 of them were monomeric repeats (Table 1). In particular, 81.3% of the total SSRs were found in LSC, and dimeric and trimeric repeats were also found in the LSC region. Nine monomeric repeats were found in the SSC region, and the five were in the IRs region.

Results of phylogenetic analysis showed that the genomes of *Hibiscus* formed a well-supported clade (Fig. 2). Among the seven *Hibiscus* species in the phylogenetic tree, *H. sabdariffa* is sister to all other species of *Hibiscus*. The complete chloroplast genomes of *Hibiscus* are too few to correctly infer the phylogenetic relationships of the members of the genus, considering the number of species of the genus. However, our phylogenetic analysis suggests that *H. sabdariffa* is distinct. The complete chloroplast genome of *H. sabdariffa* determined in this stduy will provide useful information for further studies on evolution and biodiversity with *Hibiscus* species and Malvalceae, which contain many economically important plants.

ORCID: Soon-Ho KWON 0000-0003-4142-9973; Yunmi PARK 0000-0003-2428-4581; You Lim JANG 0000-0002-6425-2448; Hae-Yun KWON 0000-0001-8365-8788.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Forest Science (NiFoS) grant funded by the project number FG0403-2018-01.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest.

LITERATURE CITED

Abdelhafez, O. H., E. M. Othman, J. R. Fahim, S. Y. Desoukey, S. M. Pimentel-Elardo, J. R. Nodwell, T. Schirmeister, A. Tawfike and U. R. Abdelmohsen. 2020. Metabolomics analysis and biological investigation of three Malvaceae plants. Phytochemical Analysis 31: 204–214.

Abdullah, F. Mehmood, I. Shahzadi, S. Waseem, B. Mirza, I. Ahmed and M. T. Waheed. 2020. Chloroplast genome of *Hibiscus rosa-sinensis* (Malvaceae): Comparative analyses and identification of mutational hotspots. Genomics 112: 581–591.

Beier, S., T. Thiel, T. Münch, U. Scholz and M. Mascher. 2017.

MISA-web: A web server for microsatellite prediction. Bioinformatics 33: 2583–2585.

Bule, M., A. H. Albelbeisi, S. Nikfar, M. Amini and M. Abdollahi. 2020. The antidiabetic and antilipidemic effects of *Hibiscus sabdariffa*: A systematic review and meta-analysis of randomized clinical trials. Food Research International 130: 108980.

Cheng, Y., L. Zhang, J. Qi and L. Zhang. 2020. Complete chloroplast genome sequence of *Hibiscus cannabinus* and comparative analysis of the Malvaceae family. Frontiers in Genetics 11: 227.

Christenhusz, M. J. M. and J. W. Byng. 2016. The number of known plants species in the world and its annual increase. Phytotaxa 261: 201–217.

Da-Costa-Rocha, I., B. Bonnlaender, H. Sievers, I. Pischel and M. Heinrich. 2014. *Hibiscus sabdariffa* L.: A phytochemical and pharmacological review. Food Chemistry 165: 424–443.

Dhar, P., C. S. Kar, D. Ojha, S. K. Pandey and J. Mitra. 2015. Chemistry, phytotechnology, pharmacology and nutraceutical functions of kenaf (*Hibiscus cannabinus* L.) and roselle (*Hibiscus sabdariffa* L.) seed oil: An overview. Industrial Crops and Products 77: 323–332.

Dierckxsens, N., P. Mardulyn and G. Smits. 2017. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. Nucleic Acids Research 45: e18.

Greiner, S., P. Lehwark and R. Bock. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Research 47: W59–W64.

Izquierdo-Vega, J. A., D. A. Arteaga-Badillo, M. Sánchez-Gutiérrez, J. A. Morales-González, N. Vargas-Mendoza, C. A. Gómez-Aldapa, J. Castro-Rosas, L. Delgado-Olivares, E. Madrigal-Bujaidar and E. Madrigal-Santillán. 2020. Organic acids from roselle (*Hibiscus sabdariffa* L.): A brief review of its pharmacological effects. Biomedicines 8: 100.

Kim, S.-T., S.-H. Oh and J. Park. 2021. The complete chloroplast genome of *Diarthron linifolium* (Thymelaeaceae), a species found on a limestone outcrop in eastern Asia. Korean Journal of Plant Taxonomy 51: 345–352.

Kwon, H.-Y., J.-H. Kim, S.-H. Kim, J.-M. Park and H. Lee. 2016. The complete chloroplast genome sequence of *Hibiscus syriacus*. Mitochondrial DNA Part A, DNA Mapping, Sequencing, and Analysis 27: 3668–3669.

Liu, X.-F., G.-F. Zhu, D.-M. Li and X.-J. Wang. 2019. Complete chloroplast genome sequence and phylogenetic analysis of *Spathiphyllum* 'Parrish'. PLoS ONE 14: e0224038.

Mandaji, C. M., R. da Silva Pena and R. C. Chiste. 2022. Encapsulation of bioactive compounds extracted from plants of genus *Hibiscus*: A review of selected techniques and applications. Food Research International 151: 110820.

Montalvo-González, E., Z. Villagrán, S. González-Torres, L. E. Iñiguez-Muñoz, M. A. Isiordia-Espinoza, J. M. Ruvalcaba-Gómez, R. I. Arteaga-Garibay, J. L. Acosta, N. González-Silva and L. M. Anaya-Esparza. 2022. Physiological effects and human health benefits of *Hibiscus sabdariffa*: A review of clinical trials. Pharmaceuticals 15: 464.

Morton, J. F. 1987. Fruits of Warm Climates. Creative Resource Systems, Inc., Winterville, NC. Pp. 281–286.

Sánchez-Mendoza, J., A. Domínguez-López, S. Navarro-Galindo and J. A. López-Sandoval. 2008. Some physical properties of roselle (*Hibiscus sabdariffa* L.) seeds as a function of moisture content. Journal of Food Engineering 87: 391–397.

Sievers, F., A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson and D. G. Higgins. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology 7: 539.

Tamura, K., G. Stecher and S. Kumar. 2021. MEGA11: Molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution 38: 3022–3027.

Tillich, M., P. Lehwark, T. Pellizzer, E. S. Ulbricht-Jones, A. Fischer, R. Bock and S. Greiner. 2017. GeSeq: Versatile and accurate annotation of organelle genomes. Nucleic Acids Research 45: W6–W11.