Chloroplast genome of the conserved Aster altaicus var. uchiyamae B2015-0044 as genetic barcode

Minjee Lee¹, Jae-Sun Yi², Jihye Park¹ and Jungho Lee¹,*

¹Green Plant Institute, B-301, Heungdeok IT Valley, Yongin 16954, Republic of Korea
²Shingu Botanical Garden, Seongnam 13443, Republic of Korea

*Correspondent: jlee@greenplant.re.kr

An endemic endangered species, Aster altaicus var. uchiyamae (Danyang aster) B2015-0044, is cultivated at the Shingu Botanical Garden, which serves as the ex situ conservation institution for this species. In this work, we sequenced the chloroplast genome of A. altaicus var. uchiyamae B2015-0044. We found that the chloroplast (cp) genome of B2015-0044 was 152,457 base pairs (bps) in size: 84,247 bps of large single copy regions (LSC), 25,007 bps of inverted repeats (IRs), and 18,196 bps of small single copy regions. The B2015-0044 cp genome contains 79 protein-coding genes (PCGs), 4 RNA genes, 29 tRNA genes, and 3 pseudogenes. These results were identical to a previously reported cp genome (Park et al., 2017), except for two sites in introns and three in intergenic spacer (IGS) regions. For the intronic differences, we found that clpP31 had a 1-bp small simple repeat (SSR) (T) and petD.i had a 3-bp SSR (ATT). We found 1-bp SSRs in the IGSs of trnT_ggu~psbD and psbZ~trnG_gcc, C and A, respectively. The IGS of (ndhF)~rpl32 had a SNP. Based on our results, the cp genome of the A. altaicus var. uchiyamae can be classified into two genotypes, [C]¹-[A]¹²-[T]¹²-[ATT]²-C and [C]²-[A]¹¹-[T]¹¹-[ATT]²-A.

Keywords: Aster altaicus, chloroplast genome, conservation, genotype, SNP, SSR

© 2021 National Institute of Biological Resources
DOI: 10.12651/JSR.2021.10.2.154

INTRODUCTION

Endangered species conservation is a hot issue in botany. In Korea, ex situ Conservation Institutions play major roles in the propagation of and finding new habitats for rare species. Aster altaicus var. uchiyamae, an endemic endangered species, has potential horticultural and agricultural uses, suggesting that there are strong merits to conserving this species. The Danyang aster (Aster altaicus var. uchiyamae) is a useful garden plant with pink ray flowers. Recently, Mr. Lee of the Shingu Botanical Garden has developed a Danyang aster with white flowers. Unlike the other Dendranthema asterid species, which are propagated by cutting (Park et al., 2015), A. altaicus var. uchiyamae is a biennial and propagated by seeds.

Aster altaicus var. uchiyamae Kitam. was first collected by K. Uchiyama on October 7, 1902 in Suanbo, Chungcheongbuk Province (Kitamura, 1937). The plant grows in limited quantities at roadsides and sandy riverbanks along the South Han River (personal observation). In the last decade, much of this species’ habitat was submerged by dam construction, thus its protection is urgent. Aster altaicus var. uchiyamae is protected by Korean law as an endangered species and is included in the Korean Red List of Threatened Species (Suh and Kim, 2014). Lines of A. altaicus var. uchiyamae have been propagated by the Korean Association for ex situ Conservation Institution (KAECI), supported by the Ministry of Environment, Republic of Korea.

Genetic information for native plants of different countries is provided in the Access to Genetic Resources and Benefit-Sharing development (ABS) under the Convention on Biological Diversity (CBD). Chloroplast (cp) genetic information has been used for more than three decades because of its maternal inheritance and sequence stability. In Aster, the first plastid genome was documented in A. spathulifolius (Choi and Park, 2015), which grows at seashores of Korea and Japan. The cp genome of A. altaicus var. uchiyamae was documented and the cp-DNA was compared to that of A. spathulifolius by Park et al. (2017). In the study, we reported that A. altaicus var. uchiyamae cp genome contains 112 genes and 21 introns and consisted of 79 protein-coding genes (PCGs), 4 RNA genes, 29 tRNA genes, and 3 pseudo-genes. Pseudo-genes include
psi-ycf1, psi-rps19, and psi-trnT_GU.

The reported plastid genome of A. altaicus var. uchiyamae (Park et al., 2017) was studied from plant material collected in 2010 before dam construction from the population at Jocheon-ri, Chungju city of Chungcheongbuk (CB) Province. Here, we characterize the cp genome of the species from a different collection site, Gyeonggi (GG) Province, to investigate genetic diversity in the cp genome. Complete cp-genomic sequences have been utilized as genetic barcodes in plants (Nock et al., 2011; Li et al., 2015). Here, we report the complete cp genome of A. altaicus var. uchiyamae as a genetic barcode.

**MATERIALS AND METHODS**

**Chloroplast DNA extraction, genome sequencing, assembly, and PCR-based validation**

*Aster altaicus var. uchiyamae* was collected under the sample name B2015-0044 as an F4 generation from seeds collected from the Gangcheon Island of Yeouju City, Gyeonggi Province, before the island was submerged by dam construction. Fresh leaves of *A. altaicus var. uchiyamae* were collected from the Shingu Botanical Garden and stored in liquid nitrogen until DNA extraction. Total DNA was extracted using the Qiagen DNeasy Plant Mini Kit stored in liquid nitrogen until DNA extraction. DNA concentration and quality were determined using a Scandrop Nano-volume spectrophotometer (Qiagen, Germany). DNA concentration and quality were determined using a Scandrop Nano-volume spectrophotometer (Qiagen, Germany). High quality DNA (concentration = 300 ng/µL, A260/280 ratio = 1.8–2.0, and A260/230 ratio = 1.7) was used for polymerase chain reaction (PCR) and sequencing.

The Illumina paired-end genomic library of 200 bps was constructed and sequenced using an Illumina HiSeq 2000 platform. The plastid sequence was obtained using CLC Genomics Workbench version 7.05 as described by Jeong et al. (2014). Circular structures of each replicon were confirmed by PCR amplification at their ends and by the joining of Sanger sequencing reads derived from the amplicons. The assemblies were further verified by examining paired-end distance and depth after re-mapping reads on the contig sequences. The BLAST searches of a large contig were verified to be plastid genomes.

**Genome annotation, Genome comparison, and Sequence Analysis**

For gene annotation of organelle genomes, protein-coding and ribosomal RNA genes were annotated using DOGMA (http://dogma.ccbb.utexas.edu/; Wyman et al., 2004). The boundaries of each annotated gene were manually determined by comparing with orthologous genes from other known cp genomes. Genes encoding tRNAs were first predicted using tRNA scan-SE; Lowe and Eddy, 1997) and ARAGORN version 1.2 (http://130.235.46.10/ARAGORN/; Laslett and Canback, 2004), and were manually verified by predicting the tRNA secondary structure. Circular genome maps were drawn using GenomeVx (Conant and Wolfe, 2008), followed by manual modification. The sequencing data and gene annotation were submitted to GenBank with accession number MK860967.

**RESULTS AND DISCUSSION**

We performed a detailed comparison of two cp genomes of *A. altaicus var. uchiyamae*, such that the Sanger-sequenced KX35265 (CB) was revised at four sites and the B2015-0044 cp genome (GG) was determined (Fig. 1). A total of 112 genes, including 79 protein-coding genes, 29 tRNA genes, and 4 rRNA genes reported in Park et al. (2017), were identical in DNA sequence. Both cp genomes of *A. altaicus var. uchiyamae* contained identical copies of psi-ycf1 and psi-rps19 in the borders of inverted repeat (IR) and single copy regions. In addition, psi-trnT_GUU copies were identical in length and sequence. The cp genomes were 152,450 bps and 152,457 bps in length (Table 1). The length was variable in large single copy (LSC) regions of 82,240 (CB population) and 82,247 (GG population) bps. However, the lengths of small single copy (SSC) region and IRs were identical, 18,196 bps and 25,007 bps (Table 1).

<table>
<thead>
<tr>
<th>Species</th>
<th>Genome size (bp)</th>
<th>LSC size (bp)</th>
<th>IR size (bp)</th>
<th>SSC size (bp)</th>
<th>Genbank ACC #</th>
<th>Coll. Site/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Aster altaicus var. uchiyamae</em></td>
<td>152,457</td>
<td>84,247</td>
<td>25,007</td>
<td>18,196</td>
<td>MK860967</td>
<td>Yeouju, Gyeonggi [GG], This study</td>
</tr>
<tr>
<td><em>Aster indicae</em></td>
<td>152,450</td>
<td>84,240</td>
<td>25,007</td>
<td>18,196</td>
<td>NC_034996</td>
<td>Cheongju, Chumbuk [CB], Park et al., 2017</td>
</tr>
<tr>
<td><em>Aster scaber</em></td>
<td>152,885</td>
<td>84,610</td>
<td>25,003</td>
<td>18,269</td>
<td>NC_040126</td>
<td>Liu et al. 2018</td>
</tr>
<tr>
<td><em>Aster spathulifolius</em></td>
<td>152,780</td>
<td>84,521</td>
<td>24,992</td>
<td>18,275</td>
<td>GPI-ASTR01</td>
<td>Jeju Island, Lee et al. unpublished</td>
</tr>
</tbody>
</table>

Dokdo, KyungBuk, Choi and Park, 2015
There are five variable sites in the cp genomes (Table 2). The detailed sequence alignment of the five variable sites is shown in Figure 2. IRs of both genomes are identical in DNA sequence, but SSCs of the two genomes have one SNP in the IGS between \textit{ndhF} and \textit{rpl32}. As well as the single SNP in SSC, small simple repeats (SSR) were found at four variable sites in LSC. The four SSR occur in two introns and two IGSs (Table 2). The introns include the first intron of \textit{clpP}, \textit{clpP.i1}, and the intron of \textit{petD}, \textit{petD.i}. The IGSs include \textit{trnT\_ggu}~\textit{psbD} and \textit{psbZ}~\textit{trnG\_gcc}.

### Table 2. Variable sites between two cp-genomes of \\textit{Aster altaicus var. uchiyamae}.

<table>
<thead>
<tr>
<th>Region</th>
<th>Site</th>
<th>Size (GG/CB)</th>
<th>Type of variation</th>
<th>Number of variation</th>
<th>Memo [GG/CB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intron</td>
<td>LSC</td>
<td>clpP_intron1</td>
<td>813/812</td>
<td>INDEL</td>
<td>1 bp SSR {T}^{11}/(T)_{11} }</td>
</tr>
<tr>
<td></td>
<td></td>
<td>petD_intron</td>
<td>766/760</td>
<td>INDEL</td>
<td>3 bp SSR { ATT}^{3}/(ATT)_{3} }</td>
</tr>
<tr>
<td>IGS</td>
<td>LSC</td>
<td>\textit{trnT_ggu}~\textit{psbD}</td>
<td>1396/1397</td>
<td>INDEL</td>
<td>1 bp SSR {C}/(C)_{7} }</td>
</tr>
<tr>
<td></td>
<td></td>
<td>\textit{psbZ}~\textit{trnG_gcc}</td>
<td>322/321</td>
<td>INDEL</td>
<td>1 bp SSR {A}^{11}/(A)_{11} }</td>
</tr>
<tr>
<td>SSC</td>
<td></td>
<td>(\textit{ndhF})~\textit{rpl32}</td>
<td>1037</td>
<td>SNP</td>
<td>C/A</td>
</tr>
</tbody>
</table>

Fig. 1. Plastid genomic map of \\textit{Aster altaicus var. uchiyamae}.
Z~trnG_gcc. A 1-bp SSR of [T]^{12} and [T]^{11} occurs in clpP.i1, and 3-bp SSRs of [ATT]^4 and [ATT]^2 occur in petD.i. An IGS of trnT_ggu~psbD has [C] and [C]^2, and IGS of psbZ~trnG_gcc has [A]^12 and [A]^{11}. The position of the variable sites in chloroplast genome structure is shown in Figure 3.

Based on these variations, the chloroplast genome of A. altaicus var. uchiyamae can be simplified as two genotypes. One genotype, GG, is [C]^{1}-[A]^{12}-[T]^{12}-[ATT]^4-C and the other genotype, CB, is [C]^2-[A]^{11}-[T]^{11}-[ATT]^2-A. The relative positions of the variable sites in the cp-map are shown in Figure 3. Though a detailed characterization of Aster cp-DNA is currently in process, it is clear that four rRNA genes (16S, 23S, 4.5S, 5S) are identical in three species of Aster (Table 1).

All 112 gene sequences and 3 pseudo-genes were identical in plant materials from two population localities that were 25 km apart. We found minor variation in five sites in introns and IGS regions, which indicates that the species is stable with regard to cp genome evolution.

**Conclusion**

We investigated two cp genomes of A. altaicus var. uchiyamae, an endangered species. One genome (CB, KX35265) was from the extinguished population of Chungcheongbuk Province. The other genome (GG, MK860967) was the conserved line B2015-0044 from the Shingu Botanical Garden, which was propagated by seeds from Gyeonggi Province. The chloroplast genomes were 152,450 and 152,457 bps, respectively. GG cp-DNA contains 79 protein-coding genes (PCGs), 4 RNA genes, 29 tRNA genes, and 3 pseudo-genes that were identical to CB cp-DNA. We found differences in five sites at introns and IGS. The introns are clpP.i1 and petD.i, which had a 1-bp SSR (T) and a 3-bp SSR (ATT), respectively. The IGSs include trnT_ggu~psbD, psbZ~trnG_gcc, and (ndhF)~rpl32. IGSs of trnT_ggu~psbD and psbZ~trnG_gcc have 1-bp SSRs, C and A, respectively. The IGS of (ndhF)~rpl32 has a SNP. Based on results, two genotypes of Danyang aster were identified, GG type [C]^{1}-[A]^{12}-[T]^{12}-
[ATT]^{2-C} and CB type [C]^{2-[A]}^{11-[T]}^{11-[ATT]}^{2-A}. The descendants of B2015-0044 at the Shingu Botanical Garden can be classified as the GG type [C]^{1-[A]}^{12-[T]}^{12-[ATT]}^{4-C}. Additionally, more lines of *A. altaicus* var. *uchiyamae* have been propagated in other *ex situ* conservation institutions. The characterization of these additional lines of the species would be useful for tracing the origin of their respective habitats and the *ex situ* conservation institution, before the species are propagated by the general public.

**ACKNOWLEDGEMENTS**

This research was funded by the National Research Foundation of Korea (NRF), grant number NRF-2019R1F1A1062617, and was also supported by a grant from the National Institute of Biological Resources under the Ministry of Environment (NIBR201703201), Republic of Korea.

**REFERENCES**


Submitted: January 3, 2021
Revised: February 2, 2021
Accepted: February 3, 2021