Research & Reviews: Research Journal of Biology

e-ISSN:2322-0066

Comparative Chloroplast Genome Analysis of Suaeda aralocaspica with Other Amaranthaceae Genomes

Caburatan L , Park J*

Department of Fine Chemistry, Seoul National University of Science and Technology, Seoul, South Korea

Review Article

Received date: 02/10/2019

Accepted date: 12/10/2019

Published date: 21/10/2019

***For Correspondence**

Joonho Park, Department of Fine Chemistry, Seoul National University of Science and Technology, Seoul, South Korea.

E-mail: jhpark21@seoultech.ac.kr

Keywords: *Suaeda aralocaspica, chloroplast genome, sequence analysis, tandem repeats, IRa, IRb*

ABSTRACT

Chloroplasts are semi-autonomous organelles needful for the food manufacturing system of plants. Chloroplast genomes are widely studied nowadays for comparative analysis which may help elucidate the differences between species of the same plant family. *Suaeda aralocaspica* is a single-cell C4 plant species belonging to the family Amaranthaceae under subfamily Suaedoideae. However, not all plant species under Suaedoideae utilize C4 photosynthesis as most of its plant species utilize C3 photosynthesis. This study therefore aims to evaluate the chloroplast genome of *S. aralocaspica* and conduct a comparative analysis with four other Amaranthaceae species. Chloroplast genome of *S. aralocaspica* was obtained by chloroplast isolation and DNA sequencing as well as by genome assembly and annotation. Genome comparison was conducted by sequence analysis of genes as well as the determination of tandem repeats. *S. aralocaspica* is 151,225 bp length with a coding size of 77,517 bp. It has a total of 140 genes, 86 of which are protein-coding. There are 20 duplicated genes in both IRa and IRb and there are 16 intron-containing genes. Results from 70% cut off sequence identity showed that *S. aralocaspica* varied with other species more apparently in the arrangement of few genes such as matK, atpF, rpoC2, rpoC1, ycf6, psbM, clpP, petB, rpl16, ndhA, and ycf1. A total of 66 tandem repeats were found in all of the studied species, most of which is situated in the intergenic space. Comparative chloroplast genome analysis showed that *S. aralocaspica* is much closer to *Suaeda glauca* and a little in variance with *Suaeda acuminata* and the outgroup, *Salsola tragus*.

INTRODUCTION

Chloroplasts are semi-autonomous organelles essential for the production of solar energy into carbohydrates [1] for the food manufacturing system of plants. Each leaf of plants approximately contains 1,000-10,000 chloroplasts [2]. These organelles also play key roles in other biochemical pathways such as the biosynthesis of starch, pigments, fatty acids, and amino acids [3]. Chloroplasts are also known to have their own genome and thus, are considered to contain their own DNA and gene expression systems [4]. These circular genomes are about 120-160 kb in size [5] and are highly conserved in higher plants [3].

Plants in the family Amaranthaceae comprises about 70 genera and 1000 species. Amaranthaceae species are able to tolerate highly arid habitats and very salty soils [6] and can either be utilizing the C3 or C4 photosynthetic pathway. Among the plants present under this family belong to the subfamily Suaedoidae and Salsoloidaea. In subfamily Suaedoideae, the genus Suaeda contains 40 C4 species out of about 82 spp [7].

In this study, *Suaeda aralocaspica* chloroplast genome was elucidated to further understand its genomic structure in relation to other Amaranthaceae species. *S. aralocaspica* like most plants in the genus Suaeda is a halophyte which utilizes a C4 photosynthesis and is considered as one of the unique single-cell C4 (SCC4) species. SCC4 species do not possess mesophyll cells (MC) and bundle sheath cells (BSC) but rather have two biochemically and morphologically different chloroplast types, spatially separated between two cytoplasmic domains within individual photosynthetic cells. These two chloroplast types in *Suaeda aralocaspica* are arranged in elongated chlorenchyma cells, proximal and distal with respect to the internally located veins and water storage tissue [8].

The efficiency of the C4 photosynthesis to suppress photorespiration and increase carbon gain in response to high light intensities and high temperature environment make it a center of interest for many studies. In addition to this, there also is a

growing interest to study about SCC4 species due to its unique chloroplast organization. In this study, information about plastid genome of a SCC4 species can be a good resource for phylogenetic analysis and DNA barcoding and may also be considered as a basic tool for future chloroplast engineering prospects.

MATERIALS AND METHODS

Genome sequencing preparation

Plants of *Suaeda aralocaspica* were grown in a growth chamber at 26°C under a 16 h light and 8 h dark cycle. Fresh leaves (100 g) of *S. aralocaspica* were collected. Chloroplast DNA was extracted following the protocol of WizPrepTM Plant DNA mini kit (Wizbiosolutions, South Korea) and the total DNA concentration was measured using Optizen POP UV/Vis spectrophotometer (Mecasys, South Korea) at 260 nm. The chloroplast genome of *S. aralocaspica* was analyzed using a combined approach with 454 GS FLX Titanium system (Roche Diagnostics, Brandford, CT) with an 8-kb paired-end library and the Illumina GAIIx (San Diego, CA). The 454 GS FLX sequencing achieved about 3.5-fold coverage, while 290.2-fold read coverage was achieved by Illumina paired-end sequencing. The reads generated by the Illumina GAIIx and the 454 GS FLX Titanium were assembled using Celera Assembler 7.0.

Genome assembly and annotation

De novo assembly was conducted using Celera Assembler 7.0 [9]. Gene prediction and annotation were carried out using Glimmer3, the RAST annotation server, and the NCBI COG database. Geneious version 8.1.6 was used to annotate the chloroplast genome and manual evaluation of annotation results was conducted. Homologue genes from previously known chloroplast DNA was used to compare codon positions with slight adjustments. The circular DNA map was drawn through OGDRAW software [10]. The chloroplast genome of *S. aralocaspica* was deposited to GenBank in forms of Sequence Read Archive (SRA) with accession number: PRJNA491597.

Genome comparison with other Amaranthaceae species and sequence analysis

S. aralocaspica chloroplast genome was obtained using the software Geneious v. 8.1.6 (Biomatters, NZ). Genomes of other Amaranthaceae species were obtained from GenBank with the following accession number: *Suaeda glauca*, PRJNA497752; *Suaeda acuminata*, PRJNA497764; *Salsola tragus*, PRJNA497767; and, *Bienertia sinuspersici*, KU726550 [11]. Aside from *B. sinuspersici* chloroplast genome, all other genomes were sequenced and loaded at NCBI through SRAs.

Genome comparison of chloroplast genome with the above mentioned species was conducted through an evaluation of genome size, gene loss, and evaluation of the sequential analysis of genes. Sequence analysis of genes was performed using the mVISTA program in Shuffle-LAGAN mode with 70% cut-off identity. The genome sequence of *S. aralocaspica* was used as the reference, while the outgroup was that of *Salsola tragus*. Tandem repeats were also determined using Tandem Repeats Finder version 4.09 [12].

RESULTS AND DISCUSSION

Chloroplast genome features of S. aralocaspica

Chloroplast genome of *S. aralocaspica* is a double stranded circular structure with 151,225 bp length. Its quadripartite structure is subdivided into a large single copy (LSC) of 82,296 bp and small single copy (SSC) of 17,939 bp separated by two inverted repeat (IR) regions of 50,346 bp each.

The positions of the genes present in the chloroplast genome of *S. aralocaspica* is shown in Figure 1 along with its corresponding functional categorization. There are 140 predicted functional genes, 120 of which are unique composing of 86 protein-coding genes, 30 tRNAs, and 4 RNA genes. The LSC region contains 84 genes while 13 genes comprise the SSC. Duplicated genes in the IR regions amounted to 20, most of which are protein-coding genes that includes 3 hypothetical proteins. Gene duplication is believed to contribute to genomic instability which further lead to gene rearrangement and speciation [13] and further contributes to the existence of novel functions such as the production of floral structures, induction of disease resistance, and adaptation to stress [14].

A total of 17 intron-containing genes are present in *S. aralocaspica*, 15 contain a single intron comprising of nine protein-coding and six tRNA genes. Two genes composed of clpP and ycf3 contain two introns. The shortest intron is found in rps12 which is 542 bp while the longest intron found in trnL-UAA accounts for 2,500 bp. Protein-coding genes accounted for 51.26% (77,517 bp) of the whole genome while the tRNA and rRNA genes accounted for 1.80% (2,716 bp) and 5.98% (9,047 bp), respectively. The remaining regions of the chloroplast genome were composed of non-coding sequences (CNS) which includes intergenic spacers and hypothetical proteins.

e-ISSN:2322-0066

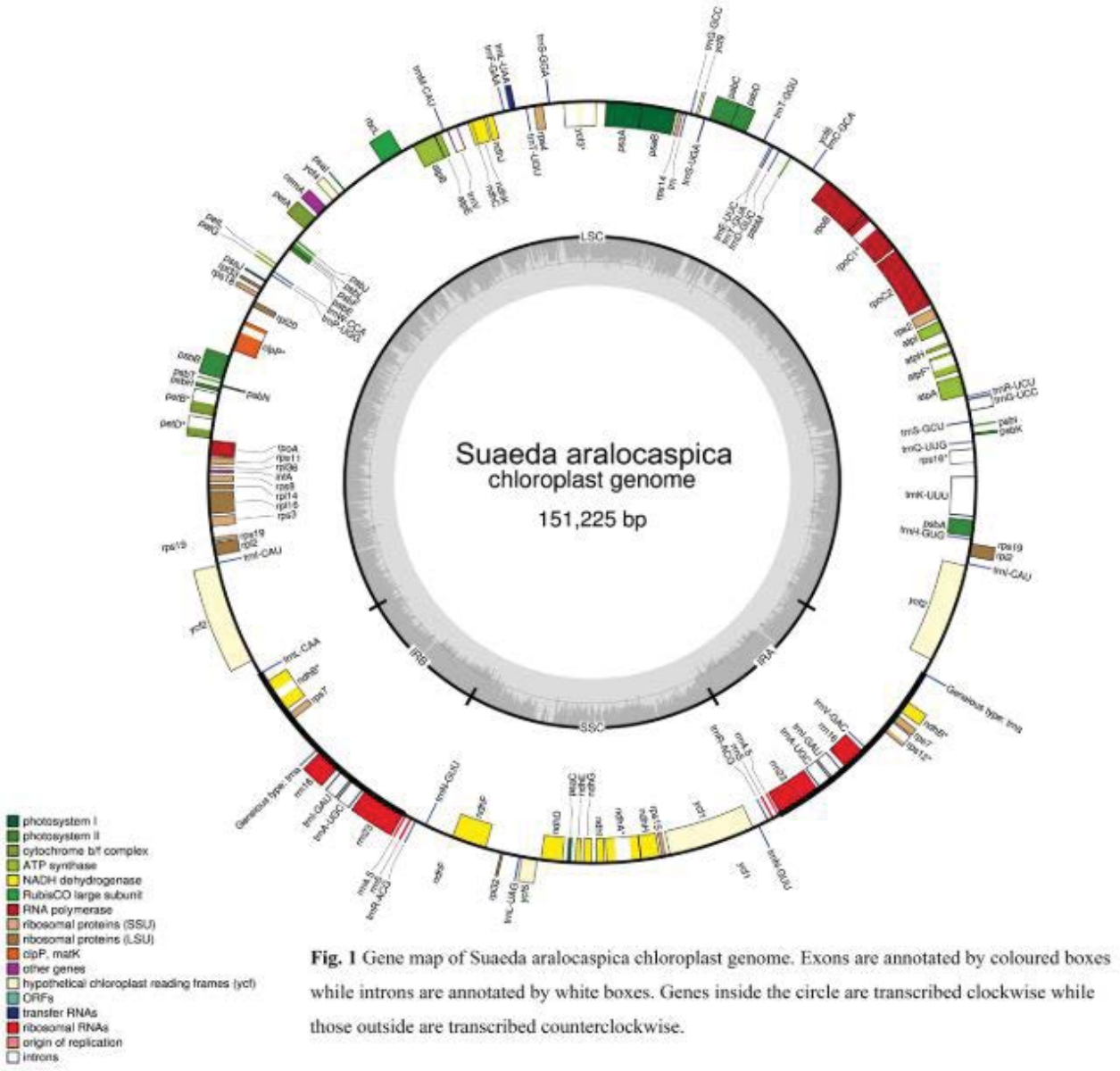

Fig. 1 Gene map of Suaeda aralocaspica chloroplast genome. Exons are annotated by coloured boxes while introns are annotated by white boxes. Genes inside the circle are transcribed clockwise while those outside are transcribed counterclockwise.

Figure 1. Gene map of Suaeda aralocaspica chloroplast genome. Exons are annotated by coloured boxes while introns are annotated by white boxes. Genes inside the circle are transcribed clockwise while those outside are transcribed counterclockwise.

Comparison with other chloroplast genomes in the family Amaranthaceae

S. aralocaspica chloroplast genome was compared to five other Amaranthaceae species. Four of the species (*Suaeda glauca*, *Suaeda acuminata*, and *Bienertia sinuspersici*) including *S. aralocaspica* belongs to the subfamily Suaedoideae while the outgroup, *Salsola tragus* belongs to the subfamily Salsoloideae. It is also interesting to note that all of these species utilize the C4 photosynthetic pathway except for *S. glauca* which is a typical C3 plant. *S. aralocaspica* and *B. sinuspersici* are especially interesting in that both are single cell C4 plants.

Notable differences observe among these species will be concentrated to genome size, gene loss, sequence divergence, and the presence of tandem repeats, as well as the expansion and contraction of the IR regions.

Genome size

The chloroplast genome size of the five species range from roughly 151 kilo base pairs up to 153 kilo bp. B. sinsuspersici has the largest genome size (153,472 bp), 2,247 bp more than *S. aralocaspica*. S. acuminata on the other hand, has the smallest genome size among the five (150,823), 402 bp smaller than *S. aralocaspica*. Table 1 summarizes the differences of general features among the species. Significant difference in the size of the LSC and the variation of the IR length as well as the length of the protein-coding region contribute to the observed differences in genome size. The outgroup, S. tragus has the largest

Table 1. General chloroplast genome features of five Amaranthaceae species.

Features	S. aralocaspica	S. glauca	S. acuminata	S. tragus	B. sinuspersici
Genome size	1,51,225	1,51,278	1,50,787	1,50,823	153472
Coding size	77,517	78,018	81,363	81,363	74,612
Spacer size	45,362	44,959	39,032	39,032	62,508
Intron size	16,583	16,537	18,556	18,582	16,352
Total gene	140	139	125	142	130
Protein-coding gene	86	89	89	90	80
Duplicated gene	20	20	21	21	17
tRNA gene	30	30	30	30	29
rRNA gene	4	4	4	4	4
Genes with intron	16	16	16	16	16

protein-coding region among the compared species, 3,856 bp higher than *S. aralocaspica*. The latter in this respect is second to the least as to the length of the coding region (77,517 bp) next to *B. sinuspersici* (74,612 bp). This result only speaks that both of these SCC4 species have the most prominent noncoding sequences (CNS) than the others.

The implication of CNS in the chloroplast genome may play significant role that may be elucidated further in the future as several studies in the present were starting to evaluate CNS significance. CNS has been known to be correlated to gene regulation and phenotype. Having this result of a higher CNS in both SCC4 species of *S. aralocaspica* and *B. sinuspersici* may also pave interest as to the urgency of a further study on CNS function and significance.

Gene loss

Loss of genes has been inferred to as one of the factors that leads to genomic differentiation not only among different plant species but also among different cultivars of the same species. In fact, in a study of Schnable and colleagues [15] in maize, they asserted that biased gene loss and expression explain at least in part, the remarkable genetic diversity found among modern maize cultivars. In this study, gene loss was also observed among the Amaranthaceae species. The genes accD and ycf5 are absent in *S. aralocaspica* and *B. sinuspersici*, respectively. Both genes has been recommended as putative plant barcodes. However, their absence in some of the species of major plants such as grass and bryophytes, respectively disqualified them as consideration as widely applicable plant barcodes [16]. In other studies, accD is also either lost such as in some cereal plastid genomes [17] or non-functional [18]. The chloroplast accD gene encodes the β-carboxyl transferase subunit of acetyl-CoA carboxylase and can be present in chloroplasts of most flowering plants, including non-photosynthetic parasitic plants and has been studied in tobacco as an essential gene required for leaf development [19].

Moreover, other genes were also found to be absent in the Amaranthaceae species under study but present in *B. sinuspersici*. These includes ccsA, psbZ, and rpl23. The plastid gene ccsA encodes a protein mediating the attachment of heme to c-type cytochromes during cytochrome biogenesis [20,21] and is localized in the plastid SSC region [22]. In other studies, this gene is also either lost [23,24] or pseudogenized [25, 26]. On the other hand, sizes of sequences of psbZ plastid genes were also recorded to be highly rearranged in ferns [27]. This gene is said to be structurally related to PSII cores in tobacco such that the psbZ-dependent interaction of PSII cores with the peripheral antenna has significant consequences for the ability of PSII to regulate the flux of light excitation [28]. The gene rpl23 is also a pseudogene in other plant species such as Cuscuta reflexa as mentioned in the study of Bommer and colleagues [29] and has been known to be a pseudogene in several angiosperm taxa [22].

Another loss of an rpl gene, rpl22 can be found in both *S. aralocaspica* and *S. glauca*. This gene has also been reported as a pseudogene in some plant species such as in Castanea mollissima and is putatively lost in some rosids as an independent transfer of rpl22 to the nucleus likely occurred [30]. The plastid gene petN is also lost in both S. acuminata and S. tragus. As the protein encoded by this gene plays significant role in photosynthesis electron transport, absence of petN may be supposed to be due to its transfer into the nuclear genome or there is a nuclear-encoded gene product that serves the same function as the petN protein as assumed in a study of the moss Tortula ruralis by Oliver and colleagues [31].

Furthermore, the loss of trnG-GCC confirms the loss of one tRNA gene in the case of *B. sinuspersici*. Other genes present in *S. aralocaspica* and the other three species and yet absent in *B. sinuspersici* includes ycf5, ycf6, and ycf9.

Sequence divergence and tandem repeats

Analysis of sequence divergence was plotted in a cut-off of 70% identity with *S. aralocaspica* as the species of reference (Figure 2). It can be noted from the presented that much divergence can be found in both LSC and SSC compared to the much reserved IR regions. The protein-coding region is also much less divergent than the non-coding region. However, some part of the

Table 2. Tandem repeat distribution among five Amaranthaceae chloroplast genome.

	Intergenice Space	Coding Sequence
S. aralocaspica	10	2
S. glauca	10	3
S. acuminate	5	3
S. tragus	3	2
B. sinuspersici	17	11
Total	45	21

coding region also contain fairly observable divergence which includes that of matK, atpF, rpoC2, rpoC1, ycf6, psbM, clpP, petB, rpl16, ndhA, and ycf1. Over-all analysis as to sequence divergence showed that *S. aralocaspica* is much similar to *S. glauca* and *B. sinuspersici* and is less similar to *S. acuminata*, much less to the outgroup, *S. tragus*.

Moreover, differences in the sequential analysis of genes can also be observed among these species. As compared to *S. aralocaspica*, an additional insertion of a hypothetical protein in the reverse direction (R) in both *S. glauca* and *S. acuminata* can be observed. The presence or absence of additional hypothetical proteins in each of the species also contribute to the variation of sequential gene arrangement as well as the loss of some genes. Differences in gene orientation can also be found such as the case of petL in *S. aralocaspica* which is oriented forward (F) as well as in all other species except in *S. acuminata*.

Tandem repeats among the five species are also present along the intergenic space (45) and in the coding sequence region (21) which sum up to 66 tandem repeats (Table 2). Most of these repeats are found in the LSC (27) and in each of the IR (18) while only 3 are found in the SSC. Repeat sequences in *S. aralocaspica* and *S. glauca* are highest in the intergenic space (10) while repeat sequences of *S. aralocaspica* in the coding region is lowest at a sum of 2 tandem repeats. *B. sinuspersici* has the highest number of tandem repeats both in the intergenic space (17) and at the coding sequence (11) while *S. tragus* has the least number of tandem repeats in the intergenic space (3) and in the coding region (2). Most repeats ranged from 30-44 bp and the longest repeat is 177 bp in *B. sinuspersici* (Figure 3a). The longest repeat in *S. aralocaspica* is around 74 bp. These repeats are more observed in the intergenic space (68%) than in the coding region (32%) as shown in Figure 3b. No repeats were found in the regions of the introns in the species under study.

Moreover, four genes are found in the coding region where tandem repeats occur: ycf2 at the IR regions of *S. aralocaspica*, *S. glauca*, and *S. acuminata*; ycf1 at the SSC of *S. glauca*; cemA at the LSC of *S. acuminata*; petB at the LSC of *S. tragus*; and, rps19 at the LSC of both *S. tragus* and *B. sinuspersici*.

IR expansion and contraction

Size variation of the chloroplast genome can be due to the expansion and contraction of its regions primarily the IRs as well as that of the SC. The SC and IR boundaries of the plastid genome of species under study is presented in Figure 4.

IR regions are said to have higher conservation of nucleotide sequences which may provide effective means of correcting inevitable mutations [32]. In a study of Palmer and Thompson [33], they found that the arrangement of nucleotide sequences in the chloroplast genomes that possess the inverted repeat structure is highly conserved, while there is scrambled arrangement of sequences in genomes that have lost the IR. This only implies that the inverted repeats have a significant role related to genomic stability. In this study, the highest IR length is found in *S. glauca* (50,402 bp), 56 bp larger than *S. aralocaspica* (50,346 bp). The smallest IR among the studied species is found in *S. tragus* (46,616 bp) which is 16 bp lesser than *S. acuminata* (46,632 bp). These differences of the IR can be attributed to the length variation of the nucleotide sequences of the duplicated genes. Both *S. aralocaspica* and *S. glauca* have 20 duplicated genes while both *S. acuminata* and *S. tragus* have 21. *B. sinuspersici* has the least number of 17 duplicated genes among the species studied.

The IRa/LSC border in all of the five species is located upstream the trnH-GUG gene. This gene is located just right at the IRa/LSC border in all of species except that of *B. sinuspersici* which is extended at a distance of 22 bp. The gene ycf1 in the IRa region expanded in *S. aralocaspica* and *S. glauca* by 1,481 bp and 1,463 bp, respectively. On the other hand, the IRa region expanded by 1,157 bp in *B. sinuspersici* as it entered the 5' end of ycf1. However, the ycf1 of both *S. acuminata* and *S. tragus* did not expand to the IRa region.

Furthermore, at the IRb/SSC junction, a hypothetical protein expands 1 bp away from the IRb region in both *S. aralocaspica* and *S. glauca* while a trnN gene sits right at the border of the IRb in both *S. acuminata* and *S. tragus*. The rps19 on the other hand lies at the border of the IRb/LSC junction in all of the five species. It expands around 143 bp to 270 bp from the border of IRb beside rpl2. The intron-containing rps12 gene in all of the species under study is transpliced with the 5' end of the gene

e-ISSN:2322-0066

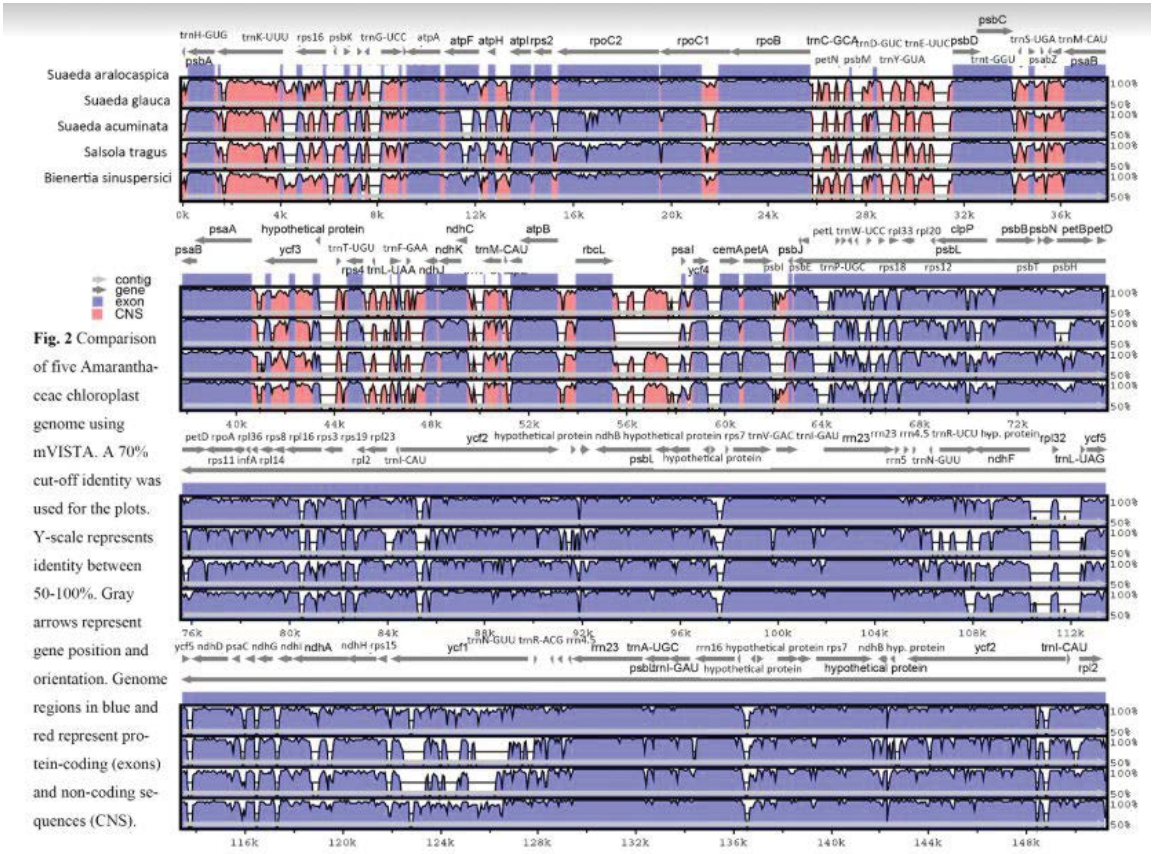

Fig. 2 Comparison of five Amaranthaceae chloroplast genome using mVISTA. A 70% cut-off identity was used for the plots. Y-scale represents identity between 50-100%. Gray arrows represent gene position and orientation. Genome regions in blue and red represent protein-coding (exons) and non-coding sequences (CNS).

Figure 2. Comparison of five amaranthaceae chloroplast genome using mVISTA. A 70% cut-off identity was used for the plots, Y-scale represents identity between 50-100%. Gray arrows represent gene position and orientation. Genome regions in blue and red represent protein-coding (exons) and non-coding sequences (CNS).

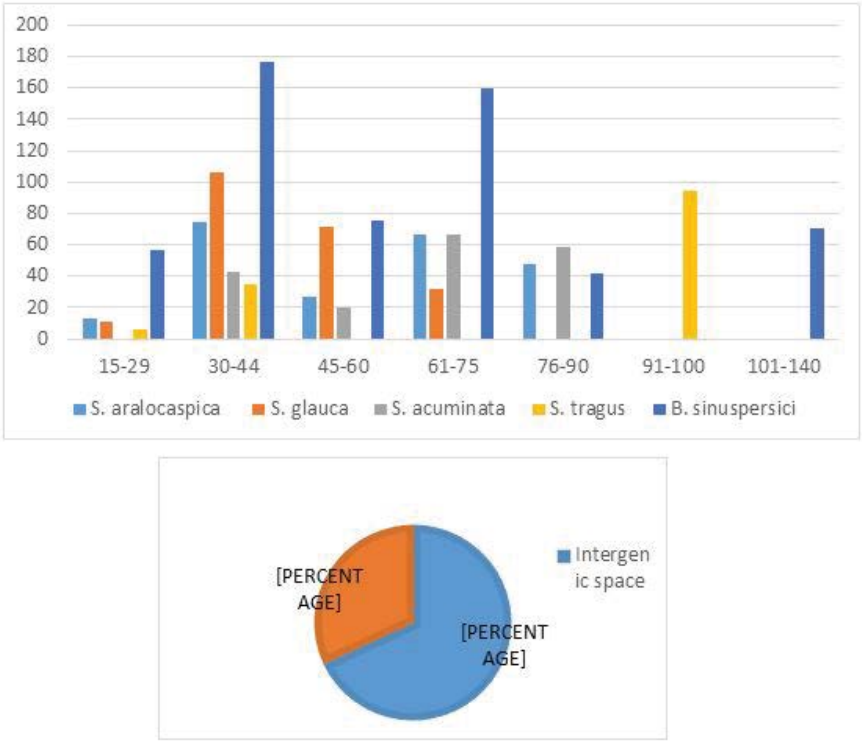

Figure 3. Gene Tandem repeat analysis in sequences of Amaranthaceae species with tandem repeat sequence length subdivided into 30 base pairs (a) and the corresponding location distribution of all tandem repeats in percent (b).

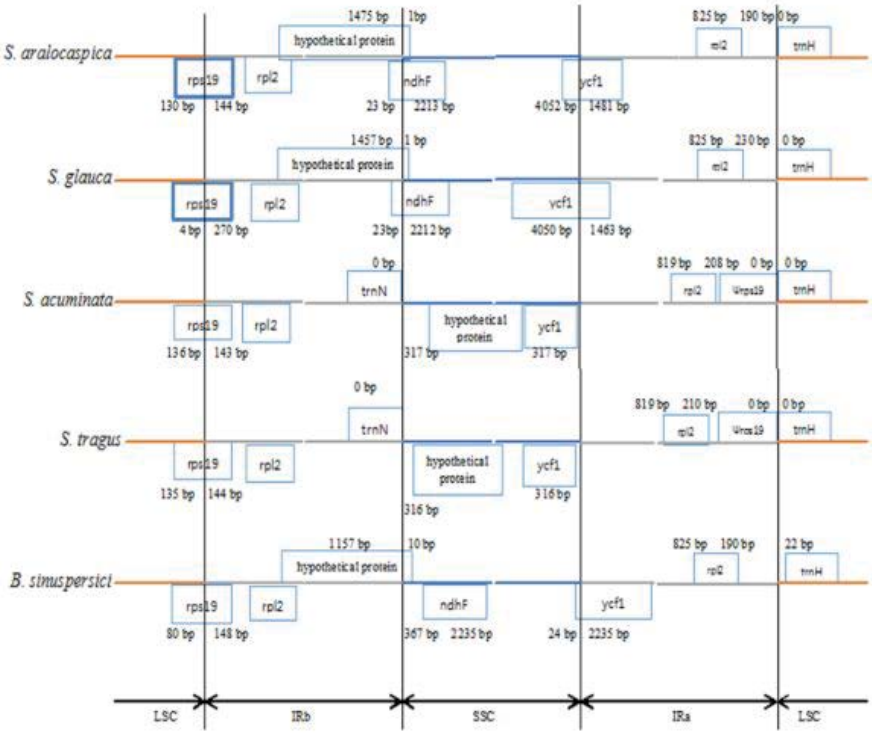

Figure 4. Comparison of the junctions between IR and single copy regions among five amaranthaceae chloroplast genome. Blue boxes represents annotated gene while the symbol for the Greek letter psi (Ψ) refers to a pseudogene.

located in the LSC and the duplicated 3' end in the IR regions. This gene is known to have vital function in translation initiation in *Chlamydomonas reinhardtii* [34].

CONCLUSION

In this study, a comparative analysis of *S. aralocaspica* chloroplast genome was conducted along with four other Amaranthaceae species from subfamily Suaedoideae except for *Salsola tragus* which was considered an outgroup as it belongs to subfamily Salsoloideae. Differences in the plastid genomes of the species under study include genome size, gene loss, sequence divergence and tandem repeats as well as the expansion and contraction of the IR regions. Results in these study showed that *S. aralocaspica* is much closer to *S. glauca* and a little in variance with S. acuminata and S. tragus based on the factors that has been mentioned above. Differences in tandem repeats can also be accounted for the plastid genome variation most of which are found in the intergenic space. Data from this paper can be helpful for any future genomic studies.

ACKNOWLEDGEMENTS

This work was carried out with the support of "Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ01095306)" Rural Development Administration, Republic of Korea.

REFERENCES

1. Yagi Y, et al. Recent advances in the study of chloroplast gene expression and its evolution. Front Plant Sci. 2014;5:61.
2. Dong W, et al. Comparative analysis of the complete chloroplast genome sequences in psammophytic Haloxylon species (Amaranthaceae). PeerJ. 2016;4:2699.
3. Raman G, et al. Analysis of the complete chloroplast genome of a medicinal plant, Dianthus superbus var. longicalyncinus, from a comparative genomics perspective. PLoS One. 2015;10:0141329.
4. Hanaoka M, et al. SIG1, a sigma factor for the chloroplast RNA polymerase, differently associates with multiple DNA regions in the chloroplast chromosomes *in vivo*. Int J Mol Sci. 2012;10:12182-12194.
5. Yang Y, et al. Complete chloroplast genome sequence of poisonous and medicinal plant Datura stramonium: organizations and implications for genetic engineering. PLoS One. 2014;11:110656.

6. Eshete MA, et al. A review on taxonomic and use diversity of the family Amaranthaceae in Ethiopia. J Med Plants Stud. 2016;2:185-194.
7. Voznesenskaya EV, et al. Unique photosynthetic phenotypes in Portulaca (Portulacaceae): C3-C4 intermediates and NAD-ME C4 species with Pilosoid-type Kranz anatomy. J Exp Bot. 2016;2:225-239.
8. Offermann S, et al. Resolving the compartmentation and function of C4 photosynthesis in the single-cell C4 species *Bienertia sinuspersici*. Plant Physiol. 2011;1:110.
9. Myers EW, et al. A whole-genome assembly of Drosophila. Science. 2000;5461:2196-2204.
10. Lohse M, et al. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr Genet. 2007;6:267-274.
11. Kim B, et al. The complete chloroplast genome sequence of *Bienertia sinuspersici*. J Mito DNA Part B 2016;1:388-389.
12. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 1999;2:573-580.
13. Cotton JA. The impact of gene duplication on human genome evolution. Els. 2008.
14. Panchy N, et al. Evolution of gene duplication in plants. Plant Physiol. 2016;4:2294-22316.
15. Schnable JC, et al. Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. Proc Natl Acad Sci. 2011,10:4069-4074.
16. Kress WJ, et al. A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. PLoS One. 2007;6:508.
17. Wakasugi T, et al. The genomics of land plant chloroplasts: gene content and alteration of genomic information by RNA editing. Photosynth Res. 2001;1:107-118.
18. Martínez-Alberola F, et al. Balanced gene losses, duplications and intensive rearrangements led to an unusual regularly sized genome in Arbutus unedo chloroplasts. PLoS One. 2013;11:79685.
19. Kode V, et al. The tobacco plastid accD gene is essential and is required for leaf development. Plant J. 2005;2:237-244.
20. Xie Z, et al. The plastid-encoded ccsA gene is required for heme attachment to chloroplast c-type cytochromes. J Biol Chem. 1996;9:4632-4639.
21. Saint-Marcoux D, et al. Biogenesis of cytochrome b6 in photosynthetic membranes. J Cell Biol. 2009;7:1195-1207.
22. Wicke S, et al. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol. 2011;3-5:273-297.
23. Wolfe KH, et al. Rapid evolution of the plastid translational apparatus in a nonphotosynthetic plant: loss or accelerated sequence evolution of tRNA and ribosomal protein genes. J Mol Evol. 1992;4:304-317.
24. Logacheva MD, et al. Sequencing and analysis of plastid genome in mycoheterotrophic orchid Neottia nidus-avis. Genome Biol Evol. 2011;3:1296-1303.
25. Wickett NJ, et al. Distribution and evolution of pseudogenes, gene losses, and a gene rearrangement in the plastid genome of the nonphotosynthetic liverwort, Aneura mirabilis (Metzgeriales, Jungermanniopsida). J Mol Evol. 2008;1:111-122.
26. Li X, et al. Complete chloroplast genome sequence of holoparasite Cistanche deserticola (Orobanchaceae) reveals gene loss and horizontal gene transfer from its host Haloxylon ammodendron (Chenopodiaceae). PloS One. 2013;3:58747.
27. Gao L, et al. Evolution of the rpoB-psbZ region in fern plastid genomes: notable structural rearrangements and highly variable intergenic spacers. BMC Plant Biol. 2011;1:64.
28. Swiatek M, et al. The chloroplast gene ycf9 encodes a photosystem II (PSII) core subunit, PsbZ, that participates in PSII supramolecular architecture. Plant Cell. 2001;6:1347-1368.
29. Bömmer D, et al. A large deletion in the plastid DNA of the holoparasitic flowering plant Cuscuta reflexa concerning two ribosomal proteins (rpl2, rpl23), one transfer RNA (trnI) and an ORF 2280 homologue. Curr Genet. 1993;1-2:171-176.
30. Jansen RK, et al. Complete plastid genome sequences of three rosids (Castanea, Prunus, Theobroma): evidence for at least two independent transfers of rpl22 to the nucleus. Mol Biol Evol. 2010;1:835-847.

e-ISSN:2322-0066

31. Oliver MJ, et al. Chloroplast genome sequence of the moss Tortula ruralis: gene content, polymorphism, and structural arrangement relative to other green plant chloroplast genomes. BMC Genomics. 2010;1:143.

32. Presting GG. Identification of conserved regions in the plastid genome: implications for DNA barcoding and biological function. Bot. 2006;9:1434-1443.

33. Palmer JD, et al. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell. 1982;2:537-550.

34. Palmer JD, et al. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell. 1982;2:537-550.