Characterization of the complete chloroplast genome of Zephyranthes phycelloides (Amaryllidaceae, tribe Hippeastreae) from Atacama region of Chile

Roberto Contreras-Díaz a,⁎, Mariana Arias-Aburto a, Liesbeth van den Brink b

a Centro Regional de Investigación y Desarrollo Sustentable de Atacama (CORDESAT), Universidad de Atacama, Av. Copayapu 485, 1530000 Copiapó, Chile
b Department of Evolution and Ecology, Plant Ecology Group, Universität Tübingen, 72076 Tübingen, Germany

A R T I C L E   I N F O

Article history:
Received 3 May 2021
Revised 13 October 2021
Accepted 14 October 2021
Available online 22 October 2021

Keywords:
Atacama Desert
Rhodophiala
Flowering desert
Chloroplast genome
Zephyranthes phycelloides
Nucleotide variability

A B S T R A C T

Sporadic rains in the Atacama Desert reveal a high biodiversity of plant species that only occur there. One of these rare species is the “Red añañuca” (Zephyranthes phycelloides), formerly known as Rhodophiala phycelloides. Many species of Zephyranthes in the Atacama Desert are dangerously threatened, due to massive extraction of bulbs and cutting of flowers. Therefore, studies of the biodiversity of these endemic species, which are essential for their conservation, should be conducted sooner rather than later. There are some chloroplast genomes available for Amaryllidaceae species; however there is no complete chloroplast genome available for any of the species of Zephyranthes subgenus Myostemma. The aim of the present work was to characterize and analyze the chloroplast of Z. phycelloides by NGS sequencing. The chloroplast genome of the Z. phycelloides consists of 158,107 bp, with typical quadripartite structures: a large single copy (LSC, 86,129 bp), a small single copy (SSC, 18,352 bp), and two inverted repeats (IR, 26,813 bp). One hundred thirty-seven genes were identified: 87 coding genes, 8 rRNA, 38 tRNA and 4 pseudogenes. The number of SSRs was 64 in Z. phycelloides and a total of 43 repeats were detected. The phylogenetic analysis of Z. phycelloides shows a distinct clade with respect to Z. mesochloa. The average nucleotide variability (Pi) between Z. phycelloides and Z. mesochloa was of 0.02000, and seven loci with high variability were identified: psbA, trnS-GCU-trnC-GCC, trnF-DNA-trnG-UAG, trnL-UAA-trnFGAA, rbcL, psbE-petL and ndhG-ndhI. The differences between the species are furthermore confirmed by the high amount of SNPs between these two species. Here, we report for the first time the complete cp genome of one species of the Zephyranthes subgenus Myostemma, which can be used for phylogenetic and population genomic studies.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Zephyranthes phycelloides or “Red añañuca” (recognizable by its red flowers) is distributed in the Atacama, Coquimbo, Metropolitan and Maule regions, and has an altitudinal range from 0 to 2200 m (Rodriguez et al., 2018). In the Atacama Desert, specifically in the Atacama Region from Chile, Z. phycelloides germinate, flower and reproduce in a short period of time due to a phenomenon called “Desierto Florido”, which is triggered by a winter precipitation greater than 15 mm, associated with the El Niño-Southern Oscillation (ENSO) weather phenomenon (Gutiérrez, 2008). Most of the plant species that emerge during these events are endemic and exclusive to the Atacama Desert (Manrique et al., 2014, Contreras et al., 2020a), and have barely or not been studied before.

Zephyranthes phycelloides (Herb.) Nic.García, formally known as Rhodophiala phycelloides (Herb.) Hunz belongs to the highly polyphyletic family Amaryllidaceae J. St.-Hil., a group of monocotyledonous, geophytic, bulbous, petaloid, cosmopolitan plants (Meerow et al., 2000). The family is composed of 1600 species of approximately 75 genera and is widely distributed in South America, the Mediterranean and South Africa (Xu & Chang, 2017). Part of the genus Rhodophiala Presl. was recently re-classified as Zephyranthes subgenus Myostemma (Salisb.) Nic.García, and it has been estimated that approximately 17 species belong to this new subgenus...
The family Amaryllidaceae is notoriously complicated in terms of diagnosability if the origin of the individual is unknown, as the family presents a high rate of both polyphyly and hybridization (García et al., 2014, 2017, 2019). The levels of polyphyly and hybridization of the family and the genus Rhodophiala Presl have been studied with karyotyping and phylogenetic approaches (Muñoz et al., 2007, Baeza et al., 2012), but there is still no consensus in its determination (Baeza et al., 2016). Comparison of the karyotypes show that Z. phycelloides, Zephyranthes bagnoldii (Herb.) Nic.García and Zephyranthes advena (Phil.) Nic.García all possess a 2n = 18 karyotype and share the same morphometry (Baeza et al., 2012). Molecular studies of internal transcribes spacer (ITS) sequences place Z. phycelloides in a closely related monophyletic group together with Z. bagnoldii, Zephyranthes montana (Herb.) Nic.García and Zephyranthes ananuca (Phil.) Nic.García (Muñoz et al., 2011). Phylogenetic and morphometric studies in species of the tribe Hippeastreae (family Amaryllidaceae), suggest natural hybridizations might have occurred between species from the Zephyranthes subg. Myostemma and Hippeastrum genera (García et al., 2014). Genetic analysis can shed a light on the genetic diversity of these endemic species, which can be used for their conservation.

Chloroplast genomes are highly valuable in taxonomy, as they are mainly maternally inherited and highly conserved (Zhang et al., 2016; Chávez-Galarza, 2021). Their slow evolution rate compared to the nuclear genome informs about molecular evolution, RNA editing, and population genetics and can solve inter-species relationships in phylogenetic studies (Zhang et al., 2018). Chloroplast genomes are available for Amaryllidaceae species (i.e. Zephyranthes mesochloa Herb. ex Lindl. (Namgung et al., 2021), Narcissus poeticus L. (Könyves et al., 2018), Lycoris longituba Y.C.Hsu & G.J.Fan (F. Zhang et al., 2019), Hippeastrum vittatum (L’Her.) Herb. (Li et al., 2020), Hippeastrum rutulum (Ker Gawl.) Herb. (Huang, 2020), and the species of the subfamily Allioideae Herbert (Namgung et al., 2021)). However, so far, there was no complete chloroplast genome available for any of the species of Zephyranthes subgenus Myostemma. We bridged this void and characterized and analyzed the chloroplast of Z. phycelloides with NGS sequencing. Our results can serve in further chloroplast analysis and phylogeny studies of the species from the Zephyranthes subgenus Myostemma genus, but also to disentangle the genetic and evolutionary complexities of the Amaryllidaceae family in future studies.

2. Materials and methods

2.1. Plant material and genomic DNA isolation

Fresh leaves of Z. phycelloides were collected near Totoral (a small town located at 27°55’15.53”S 70°56’33.67”W) (Map of the Atacama Region in Fig. 1; MINEDUC, 2021). DNA was isolated from the leaves using the modified cetyl-trimethylammonium bromide (CTAB) protocol (Contreras et al., 2020b). The DNA was quantified with a Qubit™ 3.0 fluorometer and a Qubit™ dsDNA HS Assay Kit (Life Technologies, San Diego, CA), according to the protocol provided by the manufacturer. DNA integrity was verified with an Agilent 2100 Bioanalyzer (Agilent Technologies, San Diego, CA) with an Illumina sequencing platform. DNA integrity was verified with an Agilent 2100 Bioanalyzer (Agilent Technologies, San Diego, CA) prior to NGS sequencing. Sequencing libraries were generated by a TruSeq Nano DNA LT Kit (Illumina, San Diego, CA). The final libraries were run on an Agilent 2100 Bioanalyzer to verify the fragment size distribution and concentration. Sequencing was performed at Genoma Mayor (Universidad Mayor, Chile) with an Illumina sequencing platform. Paired-end sequences of 150 bp were generated for each read (R1 and R2). The filtered reads were assembled using SPAdes 4 software version 3.13.0 (Bankevich et al., 2012), using three k-mers parameters: - k 33, 55 and 77. The chloroplast was annotated with DOGMA software (Wyman et al., 2004) and CPGAVAS2 (Shi et al., 2019), and then manually corrected. The graphical map of the chloroplast was generated using Organellar Genome DRAW (OGDRAW) (Greiner et al., 2019), and the complete nucleotide
The chloroplast structures (LSC/IR, IR/SSC) of Z. phycelloides and a total of nine closely related species in the Amaryllidaceae (Hippeastrum rutilum (Ker Gawl.) Herb., Hippeastrum vittatum (L’Hér.) Herb., Narcissus poeticus L., Lycoris sprengeri Comes ex Baker, Lycoris sanguinea Maxim., Lycoris radiata (L’Hér.) Herb., Lycoris aurea (L’Hér.) Herb., Clivia miniata (Lindl.) Regel and Zephyranthes mesochloa Herb. ex Lindl.) of the order Asparagales were visualized and compared using IRScope (Amiryousefi et al., 2018). As the subtribe Hippeastrinae is divided in two genera, Hippeastrum and Zephyranthes (García et al., 2019), we use the sequence of whole plastome (GenBank) of this subtribe for identification of simple sequence repeat (SSR). MISA software (Beier et al., 2017) was used to identify SSR in the chloroplast genome of the subtribe Hippeastrinae, with following search parameters: ≥10 repeat units for mononucleotide SSRs; ≥5 repeat units for dinucleotide SSRs; ≥4 repeat units for trinucleotide SSRs; and ≥3 repeat units each for tetra-, penta-, and hexanucleotide SSRs. To identify the tandem repeats (forward, palindromic, reverse, and complement) of these species, REPuter (Kurtz & Schleiermacher, 1999) was used. In addition, a sliding window analysis (window length: 600 bp, step size: 200 bp) was performed to assess the variability (PI) between Z. phycelloides and Z. mesochloa chloroplasts with DnaSP v. 5 software (Librado & Rozas, 2009). The complete chloroplast genome sequence of the ten species were aligned using MAFFT v7 (Katoh & Standley, 2013). The genome sequence data were analyzed using the maximum likelihood (ML) and the Bayesian inference (BI) methods. The best-fitting nucleotide substitution model of sequence evolution, model GTR + I + G, was determined using the Akaike Information Criterion (AIC) with MrModeltest v2.3 (Nylander, 2004). The ML analyses were performed using RAxML-HPC BlackBox v.8.1.24 (Stamatakis, 2014) with 1,000 bootstrap replicates; and the BI analysis was conducted using MrBayes v.3.2 (Ronquist et al., 2012) with the CIPRES Science Gateway v3.3 (Miller et al., 2010). The Markov chain Monte Carlo (MCMC) algorithm was calculated for 5,000,000 generations, and the sampling tree for every 1,000 generations. The first 25% of generations were discarded as burn-in. In the analysis, bootstrap (BS) values were estimated in the ML, and the reliability of clades in the Bayesian analysis was evaluated by means of posterior probability (PP). The trees were visualized with FigTree (Rambaut, 2012). We additionally compared the single nucleotide polymorphism (SNP) loci of the chloroplast genome of Z. phycelloides with species from order Asparagales.

Fig. 2. Circular gene map of the chloroplast genomes of Zephyranthes phycelloides. Genes were colored according to their functional group. The GC content is represented by the dashed darker grey area in the inner circle, the lighter grey area represents AT content. Small single copy (SSC), large single copy (LSC), and inverted repeats (IRA, IRB) were indicated.
137 genes contained at least one intron (Table 2). The cp genome of *Z. phycelloides* comprises 158,107 bp and its structure contains a typical quadripartite structure with two inverted repeat regions (IRs; 26,730 bp) separated by a large single copy region (LSC; 86,129 bp) and a small single copy region (SSC; 18,352 bp) (Fig. 2, Table 1). Its length and structure are similar to those of other species of the order Asparagales, who vary in the IRs between 26,730 bp and 28,610 bp, in the LSC between 86,129 bp and 86,613 bp and in the SSC between 18,121 bp and 18,541 bp (Table 1). *Z. phycelloides* had 25 bp more than *H. vittatum*, 250 bp less than *H. rutilum* (2x), and 56 SSRs (Fig. 3A). Mononucleotide repeats were the most common repeats i.e. 46 SSRs in *H. vittatum*, 43 SSRs in *Z. phycelloides*, 36 SSRs in *Z. mesochloa* and 35 SSRs in *H. rutilum* (Fig. 3A), of which the A/T repeat was most abundant (Fig. 3B) with 45 and 41 repeats in *H. vittatum* and *Z. phycelloides*, respectively. The number of dinucleotide SSRs (AG/CT and AT/AT) was similar in all of the four species. Pentanucleotide SSRs (AAACG/CGTTC) occurred in all species, except in *Z. phycelloides* (Fig. 3B), and only one hexanucleotide SSR (ACATAT/ATATGT) was observed in *Z. phycelloides*. A total of 43, 37, 37 and 38 repeats were detected in the *Z. phycelloides*, *Z. mesochloa*, *H. rutilum* and *H. vittatum* cp genomes, respectively (Fig. 4A). The repeats in *Z. phycelloides* contained 23 palindromic, 16 forward, 3 reverse and one complement repeat, while the *Z. mesochloa* cp genome contained 22 palindromic, 15 forward, zero reverse and zero complement repeats (Fig. 4A). Palindromic and forward repeats with lengths between 30 and 40 bp were most common and abundant in the four species. However, the number of palindromic and forward repeats with lengths between 30 and 40 bp was a little higher in the cp genome of Zephyranthes.
Zephyranthes than in the Hippeastrum genus. The number of palindromic and forward repeats with length between 40 and 60 bp was similar in both genera (Fig. 4BC). The reverse repeat (range of 30–40 bp) was more abundant in Z. phycelloides than in the rest of the species (Fig. 4D).

Although normally IR regions have similar lengths within the chloroplast (Gruenstaeudl & Jenke, 2020) it has been shown that IR regions can expand or contract (Ogihara et al., 2001, Odintsova & Yurina, 2007). For that reason, we compared the information of the IR-SSC and IR-LSC limits of Z. phycelloides with other species.

Fig. 3. Analysis of simple sequence repeats (SSRs) of the Z. phycelloides, Z. mesochloa, H. rutium and H. vittatum chloroplast genomes. Total numbers of SSRs of each motif unit (A) and number of SSRs detected of each motif type (B).

Fig. 4. Repeat structure analysis of the Z. phycelloides, Z. mesochloa, H. rutium and H. vittatum chloroplast genomes. Total numbers long repeat types (Palindrome, Forward, Reverse and Complement) (A), number of palindrome repeats (B), number of forward repeats (C) and number of reverse repeats (D) by length.
Fig. 5. Comparison of chloroplast genomes between the Long Single Copy (LSC), Short Single Copy (SSC) and Inverted Repeat (IRA and IRB) junction regions amongst ten species of the order Asparagales.

Fig. 6. Sliding window analysis of the whole chloroplast of *Z. phycelloides* and *Z. mesochloa*. (Window length: 600 bp, step size: 200 bp). X-axis: Nucleotide position, Y-axis: Nucleotide diversity (Pi).
from the order Asparagales. The intergenic spacers (IGS) between rpl22-rps19 genes of *Z. phycelloides*, in the junction between LSC and IRb region (JLB), are similar in size to those of *Hippeastrum* species and *Z. mesochloa*, while there is more variation when compared with other Asparagales species (Fig. 5). Likewise, the intergenic spacers (IGS) between rps19-psbA genes, located in the JLA junction, of *Z. phycelloides* and *Hippeastrum* species are similar in size, but they differ from *Z. mesochloa* and the other Asparagales species (Fig. 5). The boundaries between IRa and SSC (JSA) were located in the *ycf1* gene, and the fragment located in the IRa region was equal in size compared to *Hippeastrum* species (978 bp), whereas it differed from *Z. mesochloa* (1268 bp) and the other Asparagales species (940 bp to 2737 bp) (Fig. 5). Similarly, the *ycf1* gene spanned the IRb/SSC region, and the fragment located at the IRb region was equal in size compared to *Hippeastrum* species (978 bp), whereas it varied between *Z. mesochloa* (1267 bp) and the other Asparagales species (between 963 bp and 1061 bp) (Fig. 5). Additionally, when comparing the four *Hippeastrinae* species, only *Z. mesochloa* showed differences from the others in the JSB junction (*ndhF* gene) (Fig. 5).

The average nucleotide variability ($P_i$) between *Z. phycelloides* and *Z. mesochloa* was estimated to be 0.00492 (ranging from 0 to 0.04000) (Fig. 6). The most variable regions in the chloroplast were located in LSC and SSC regions, whereas the IR regions had a much

![Fig. 7](image_url)

**Table 3**

Number of SNPs for six chloroplast genome.

|                | *Z. phycelloides* | *Z. mesochloa* | *H. rutulum* | *H. vittatum* | *N. poeticus* | *L. aurea* |
|----------------|-------------------|----------------|--------------|---------------|---------------|------------|
| *Z. phycelloides* | –                 | –              |              |               |               |            |
| *Z. mesochloa*   | 776               | –              |              |               |               |            |
| *H. rutulum*     | 549               | 553            | –            |               |               |            |
| *H. vittatum*    | 558               | 559            | 172          |               |               |            |
| *N. poeticus*    | 2901              | 2966           | 2774         | 2763          | –             |            |
| *L. aurea*       | 1876              | 1877           | 1674         | 1669          | 2212          | –          |
lower nucleotide diversity. Seven loci with high levels of variability (Pi greater than 0.02000) were found: psbA (Pi = 0.02167), trnS^CUC-trnG^GCC (Pi = 0.02233), trnF^GUC-trnV^GUA (Pi = 0.02000), trnL^UAA-trnG^CAG (Pi = 0.02233), rbcL (Pi = 0, 02000), psbE-petl (Pi = 0.04000) and ndhG-ndhF (Pi = 0.02833) (Fig. 6).

The results of the ML and BI trees had similar topologies when we compared the whole chloroplast genome sequences of the ten species of the order Asparagales (Fig. 7). The ML phylogenetic analysis revealed four clades, one joined Z. phycelloides, Z. mesochloa, H. rutillum and H. vittatum (BP = 100), the second clade contained N. poeticus (BP = 100), the third clade was formed by Lycoris species (BP = 100) and the fourth clade was formed by the outgroup C. miniata (Fig. 7A). The BI phylogenetic analysis from the species that form the clade Hippeastrinae showed high support (PP = 1.00) and their topology was identical to ML analysis (Fig. 7B). The first clade showed two subclades where the two species of the genus Hippeastrum (H. vittatum and H. rutillum) were joined to Z. mesochloa with strong support (ML BP = 73; BI, PP = 1.00), while Z. phycelloides (Zephyranthes subg Myosotema) was separated from them with high support (BP = 100; PP = 1.00) (Fig. 7). Z. phycelloides differed in 776, 549, 558, 2501 and 1876 SNPs from the chloroplasts of Z. mesochloa, H. rutillum, H. vittatum, N. poeticus and L. aurea respectively, while the chloroplast genome substitution events between H. rutillum and H. vittatum were only 172 SNPs (Table 3).

4. Discussion

We found that the complete chloroplast genome of Z. phycelloides is highly conserved in size, number of genes and percentage of GC content in comparison to other species of the order Asparagales. Our results are consistent with other studies, which found that the chloroplast genome and the GC content, gene composition and order of genes of the Amaryllidaceae family are highly conserved (Jimenez et al., 2020, Namgung et al., 2021). We found that Z. phycelloides, and other species of the subtribe Hippeastrinae (subfamily Amaryllidoideae) that were used in this study, presented more GC content (~38%) than the species of the subfamily Aliloideae (~37.1%; Namgung et al., 2021). Interestingly, the genome size of Z. phycelloides differed more from Z. mesochloa than from H. vittatum and H. rutillum. This difference between Z. mesochloa and Z. phycelloides is due to expansion and contraction of the ycf1 and ndhF genes in the IRs regions, confirming that IRs are highly variable due to lineage-specific expansions and contractions (Zhu et al., 2016).

Zephyranthes phycelloides possesses a similar amount of genes as H. vittatum, whereas Z. mesochloa has 2 genes less (MT323232.1; Namgung et al., 2021). This variation in the number of genes could be explained by pseudogenization and gene loss that can occur in species (Petersen et al., 2015, Li et al., 2017, Mohanta et al., 2020). For example, Z. phycelloides possesses two copies of the pseudogene ycf88 that are not observed in the annotation of Z. mesochloa (Namgung et al., 2021). Pseudogenization of rps2 and gene loss of infA causes chloroplast genome fluctuations among Aliloideae species, whereas pseudogenization of ycf15 causes the fluctuations in Amaryllidaceae species (Namgung et al., 2021). However, to confirm if pseudogenization and gene loss are common in species of the genus Zephyranthes it will be necessary to analyze more chloroplasts.

A total of 56 to 67 chloroplast simple sequence repeats (cpSSRs) were found in the cp genomes of Z. phycelloides, Z. mesochloa, H. vittatum and H. rutillum (Hippeastrinae species). Our results showed variation in the number of cpSSRs, and minor differences in the mono-di-tri-tetra-penta and hexa motifs. Chloroplast simple sequence repeats (cpSSRs) are widely used to study phylogeny, but they can also be used in ecological studies or in evolutionary processes (Daniell et al., 2016). For example, 72 cpSSRs (57 were mononucleotide, 13 dinucleotide and 2 trinucleotide) were discovered in 17 cp genomes of Aliloideae (Amaryllidaceae) species, using similar SSRs search parameters as in our study (Namgung et al., 2021). A number of 44 to 54 cpSSRs were found in seven Lycoris species (Amaryllidaceae), however the SSRs search parameters were more stringent in this study (Zhang et al., 2020). Consequently, the unique genomic codpSSRs of a species can be used to analyze the genetic variation of populations and evolutionary processes of their genus.

We found the highest number of repeat elements in Z. phycelloides (43) and the lowest in H. vittatum (38). These repeat elements are important for promoting genetic rearrangements (Wu et al., 2017). Within the Hippeastrinae species the palindromic repeat was the most common one. Similar as in our species, a total between 37 and 49 repeats were found in nine cp genomes of the Allium genus, however the number of palindromic type repeats was lower (Huo et al., 2019). Another study found 21 repeats in 17 cp genomes of Aliloideae species (Namgung et al., 2021). These repeats were mainly forward repeats, whereas palindromic repeats were only found in two species (Namgung et al., 2021). The high total number of repeats as well as the high number of reverse treats found in Z. phycelloides, are important as they give insight into genome variations, rearrangements or structural expansions within and between species (Wicke et al., 2011).

The expansion and contraction of IR regions in our species are the result of modifications in the junctions of the LSC-IR-SSC regions. These expansions and contractions cause length variations in species of the order Asparagales (Wang et al., 2008), especially in the subtribe Hippeastrinae, where these regions and their connections show clear differences between Z. phycelloides and Z. mesochloa. We found a larger expansion in the ycf and ndhF genes in the junction regions in Z. mesochloa than Z. phycelloides, making the size of the chloroplast of Z. mesochloa larger than the chloroplast of Z. phycelloides. However, the junctions of Z. phycelloides and the two Hippeastrum species are very similar. I.e. in the SSC-IR border (JSB) of Z. mesochloa the ndhF and ycf1 genes overlapped, whereas these two genes were separated by a gap in Z. phycelloides, H. rutillum and H. vittatum. This corresponds to previous studies in Aliloideae species that found that the SSC-IR border (JSB) can be variable, showing either overlap or present a gap between ycf1 and ndhF (Do et al., 2020, Namgung et al., 2021). We did not find the rps19 gene within the IR-LSC borders (JLB and JLA) in any of the four Hippeastrinae species. However, Zhang et al. (2020), did detect the rps19 gene within the JLB border in two Lycoris species (Amaryllidaceae), showing a partial duplication, while five other species showed an IR expansion, and thus a complete duplication of rps19. Successive IR expansions have shown the importance of the JLA and JLB junctions for the analysis of evolutionary processes, providing clues about the origin and evolution of species (Palmer and Stein, 1986, Goulding et al., 1996).

In the present work, seven regions with high level of variability (Pi greater than 0.02) were identified: psbA, trnS^CUC-trnG^GCC, trnL^UAA-trnG^CAG, trnL^UAA-trnF^GUA, rbcL, psbE-petl and ndhG-ndhF. Chloroplast markers 3 ycf1, ndf, trnL^UAA-p^GUA and the nuclear marker ITS rDNA have proved useful in resolving part of the phylogeny of the different genera of the tribe Hippeastrae (Garcia et al., 2014). A subsequent study uncovered the information of 18 nuclear loci and 40 nearly complete chloroplast genomes for the same purpose (Garcia et al., 2017). Even though this helped to clarify some of the complexities of the family Amaryllidaceae, especially in the subtribe Hippeastrinae, the phylogeny of the genus Zephyranthes was still unclear (Garcia et al., 2019). We believe that the seven regions with high levels of variability found in this study, can be of use to resolve these uncertainties and will be useful in phylogenetic analysis of the species of the genus Zephyranthes.
The phylogenetic results ML and BI showed that Z. phycelloides represent a distinct cladeb with Z. moschola, even though both belong to the same genus. A phylogenetic analysis of near-complete plastids showed a similar separation of Z. moschola from Zephyranthes advena (Ker Gawl.) Nic. Garcia and Zephyranthes ananuca (Phil.) Nic.Garcia, two other species from Z. subg. Mysotemna (Garcia et al., 2017), however, Z. phycelloides was not included in this work, and the more than 40 nearly complete plastids are unfortunately not available for analysis. Several studies divide the Hippeastrae tribe into clades, based on chromosome number: Z. phycelloides 2n = 18, the genus Hippeastrum 2n = 22, and Z. moschola 2n = 12 (Greizerstein and Naranjo, 1987, Muñoz et al., 2011, García et al., 2017). Our results confirm the differences between Z. phycelloides and Z. moschola, which are additionally supported by SNP differences (776 SNPs). In comparison, H. vittatum and H. rutilum, showed a much lower value of SNPs (172). Allium species (Amaryllidaceae) from Central Asian species contained 451 SNPs in protein-coding genes (Yusyopov et al., 2020).

5. Conclusion

In this study, we report the complete sequence and characterization of the chloroplast genome of Zephyranthes phycelloides. Size and number of genes is conserved, similar to species from the subtribe Hippeastrae. However, Z. phycelloides and Z. moschola showed differences in genome size and slight differences in gene number. The phylogenetic analysis showed that Z. phycelloides represent a distinct cladeb with Z. moschola. The differences between the species are furthermore confirmed by the high amount of SNPs between these two species. The information of the complete chloroplast genome of Z. phycelloides shows that phylogeny of the genus Zephyranthes is still uncertain, and urgently need taxonomic studies of all the species of the genus. The results of cpSSRs and repetitive sequences can be helpful for population analysis and evolutionary studies of the genus Zephyranthes. In addition, seven highly variable regions were detected that can be used to develop useful markers for phylogenetic analysis and to distinguish between Zephyranthes species.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was financed by the Regional Innovation Fund for Regional Competitiveness (FIC Regional, 2018) of the Regional Government of Atacama, Code BIP 40013338-0. In addition, we thank the National Forestry Corporation (CONAF) of the Region of Atacama for the authorization to sample plant species of the Flowing Desert (Permit N° 106/2017 and N° 122/2019).

References

Amiryousefi, A., Hyvönen, J., & Poczaï, P. (2018). IRscope: an online program to visualize the junction sites of chloroplast genomes. Bioinformatics (Oxford, England), 34(17), 3030–3031. https://doi.org/10.1093/bioinformatics/bty220.

Baesa, C., Almendras, F., Ruiz, E., Peñalillo, P. 2012. Estudio comparativo del cariotipo en especies de Miltiada Ravenna, Phyllaea Lindl. y Rhodophiala C Prestl (Amaryllidaceae) de Chile. Revista de La Facultad de Ciencias Agrarias 44 (2), 193–205.
