The Conservation of Chloroplast Genome Structure and Improved Resolution of Infrafamilial Relationships of Crassulaceae

Hong Chang1†, Lei Zhang1†, Huanhuan Xie1, Jianquan Liu1,2, Zhenxiang Xi1* and Xiaoting Xu1*

1 Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu, China, 2 State Key Laboratory of Grassland Agro-Ecosystems, College of Life Sciences, Lanzhou University, Lanzhou, China

Crassulaceae are the largest family in the angiosperm order Saxifragales. Species of this family are characterized by succulent leaves and a unique photosynthetic pathway known as Crassulacean acid metabolism (CAM). Although the inter- and intrageneric relationships have been extensively studied over the last few decades, the infrafamilial relationships of Crassulaceae remain partially obscured. Here, we report nine newly sequenced chloroplast genomes, which comprise several key lineages of Crassulaceae. Our comparative analyses and positive selection analyses of Crassulaceae species indicate that the overall gene organization and function of the chloroplast genome are highly conserved across the family. No positively selected gene was statistically supported in Crassulaceae lineage using likelihood ratio test (LRT) based on branch-site models. Among the three subfamilies of Crassulaceae, our phylogenetic analyses of chloroplast protein-coding genes support Crassuloideae as sister to Kalanchoideae plus Sempervivoideae. Furthermore, within Sempervivoideae, our analyses unambiguously resolved five clades that are successively sister lineages, i.e., Telephium clade, Sempervivum clade, Aeonium clade, Leucosedum clade, and Acre clade. Overall, this study enhances our understanding of the infrafamilial relationships and the conservation of chloroplast genomes within Crassulaceae.

Keywords: Crassulaceae, chloroplast genome, comparative genomics, phylogenomics, infrafamilial relationships, adaptive evolution

INTRODUCTION

The angiosperm family Crassulaceae, also known as the stonecrop family, belongs to the order Saxifragales and include 34 genera and approximately 1,400 species (APG IV, 2016; Messerschmid et al., 2020), which are predominantly perennial herbs, subshrubs, or shrubs (Thiede and Eggli, 2007). The species of Crassulaceae primarily occur in (semi-)arid and mountainous habitats of the temperate and subtropical areas (van Ham and Hart, 1998), and are distributed worldwide with centers of diversity in Mexico, southern Africa, Macaronesia, and the Himalayas (Mort et al., 2001). Physiologically, Crassulaceae are characterized by the Crassulacean acid metabolism (CAM)
photosynthetic pathway (Gontcharova and Gontcharov, 2009), which achieves a higher level of water-use efficiency than either the C_3 or C_4 pathway in water-limited environment (Nobel, 1991). A recent study has identified the crown group of Crassulaceae as one of the 30 core diversification shifts across the angiosperm phylogeny (Magallón et al., 2019). This increased diversification rate may be well associated with the CAM pathway that has been recognized as an evolutionary key innovation (Pilón-Smits et al., 1996; Quezada and Gianoli, 2011; Silvestro et al., 2014).

Until now, numerous phylogenetic studies have been performed to evaluate the infrafamilial relationships of Crassulaceae with a broad taxon sampling (e.g., Berger, 1930; van Ham and Hart, 1998; Mort et al., 2001; Mayuzumi and Ohba, 2004; Gontcharova et al., 2006, 2008; Folk et al., 2019; Messerschmid et al., 2020), which have led to the establishment of three subfamilies by Thiede and Eggli (2007), i.e., Crassuloideae, Kalanchoideae, and Sempervivoideae. Although most of these studies are based on one or a few genetic loci, the monophyly of each of three subfamilies sensu Thiede and Eggli (2007) is well supported. The inter- and intrageneric relationships within each of three subfamilies have also been extensively studied using a variety of chloroplast and nuclear loci. The smallest subfamily Crassuloideae comprise approximately 200 species in a single genus, Crassula L. (Messerschmid et al., 2020), whose intragenic relationships have been recently addressed by sampling 103 species (Bryuns et al., 2019). The subfamily Kalanchoideae include approximately 240 species in four genera (Smith et al., 2019), i.e., Adromischus Lem., Cotyledon L., Kalanchoe, and Tylecodon Toelken, and the inter- and intrageneric relationships have been exclusively assessed in a few separate studies (Gehrig et al., 2001; Mort et al., 2005; Nowell, 2008). The largest and most complex subfamily Sempervivoideae contain 28 genera and over 1,000 species (Thiede and Eggli, 2007), and the inter- and intrageneric relationships have been investigated by a considerable number of studies (e.g., Mes et al., 1997; Jorgensen and Frydenberg, 1999; Mort et al., 2002; Acevedo-Rosas et al., 2004; Fairfield et al., 2004; Mayuzumi and Ohba, 2004; Gontcharova et al., 2006; Carrillo-Reyes et al., 2008, 2009; Yost et al., 2013; Zhang et al., 2014; Klein and Kaderiet, 2015; Nikulin et al., 2016; Vázquez-Cotero et al., 2017; de la Cruz-López et al., 2019). These phylogenetic studies have led to the recognition of five tribes within Sempervivoideae by Thiede and Eggli (2007), i.e., Aeonieae (Aeonium Webb and Berth., Aichryson Webb and Berth., and Monothecae Haw.), Sedeae (Afrovenella A. Berger, Dudleya Britton and Rose, Echeveria DC., Graptopetalum Rose, Lenophyllum Rose, Pachyphytum Link et al., Pistorinia DC., Prometheum H. Ohba, Rosularia Stapf, Sedella Britton and Rose, Sedum L., Thompsonella Britton and Rose, and Villadia Rose), Sempervieae (Petrodendrum Grulich and Sempervivum L.), Telephieae (Hyloclathium H. Ohba, Kungia K.T. Fu, Meterostachys Nakai, Orostachys Fisch., and Sinocrassula A. Berger), and Umbiliceae (Phedimus Raf., Pseudosedum A. Berger, Rhodiola, and Umbilicus DC.). Furthermore, based on one nuclear (ITS) and three chloroplast markers (matK, rps16, and trnL-trnF), a recent phylogenetic study of 298 Crassulaceae species has recovered six major clades of Sempervivoideae, i.e., Telephium clade, Petrodendrum clade, Sempervivum clade, Aeonium clade, Leucosedum clade, and Acre clade (Messerschmid et al., 2020). Despite the significant progress made in the tribal and generic circumscription of Crassulaceae, phylogenetic relationships among tribes and major clades remain poorly to moderately supported or sometimes contradicted, especially within Sempervivoideae (Supplementary Figure 1).

Chloroplasts are semi-autonomous replication organelles, originating in endosymbiosis between cyanobacteria and non-photosynthetic host, and play crucial roles in photosynthesis and physiology of plants (Gao et al., 2019a; Huo et al., 2019; Li et al., 2021). The chloroplast genome is mostly a typical quadripartite circular DNA genome comprising a small single-copy (SSC), a large single-copy (LSC), and two inverted repeats (IRs) (Hu et al., 2016; Chang et al., 2020). With the development of next-generation sequencing technology, chloroplast genomes have been proven to be powerful tools for resolving vague relationships of many complicated lineages due to many advantages, such as uniparental inheritance, the relatively conserved structure, low recombination and substitution rate (Goremykin et al., 2003, 2004; Twyford and Ness, 2017; Li et al., 2020a). Although the structure and sequence of plastomes are relatively conservative, variations in structure, size, and evolutionary rates of genes have been found in many studies, which in some cases signify phylogenetical information and adaptation to environment (Chumley et al., 2006; Hu et al., 2015; Piot et al., 2018; Gao et al., 2019a; Gruzdev et al., 2019; Shrestha et al., 2019; Zang et al., 2019).

Here, we sequenced fourteen fully annotated chloroplast genomes including nine species of Crassulaceae and five species of Haloragaceae. By taking advantage of the data already available, we analyzed a total of 33 chloroplast genomes and aimed to (i) assess the structural characteristics of the chloroplast genome in a comparative framework, (ii) improve the resolution of the infrafamilial relationships within Crassulaceae, and (iii) investigate the adaptive evolution by selective pressures analysis of protein-coding genes in Crassulaceae.

MATERIALS AND METHODS

Taxon Sampling, DNA Extraction, and Sequencing

The aim of our taxon sampling was to try to obtain chloroplast genome sequences for at least one representative of each well-supported major clades. The final taxon sampling contained a total of 26 Crassulaceae species (Table 1), representing all three subfamilies sensu Thiede and Eggli (2007), and all five tribes sensu Thiede and Eggli (2007) or five out of the six major clades sensu Messerschmid et al. (2020) within Sempervivoideae. Of these complete chloroplast genomes, nine were newly generated in this study, i.e., Aeonium arboreum (L.) Webb and Berthel., Cotyledon tomentosa Harv., Crassula perforata Thumb., Graptopetalum amethystinum (Rose) E. Walth, Kalanchoe fedtschenkoi Raym.-Hamet and H. Perrier, Orostachys fimbriata (Turcz.) A. Berger,
## TABLE 1 | Structural information of the chloroplast genomes of Crassulaceae and outgroups.

| Family         | Subfamily | Clade | Species                        | Genome size (bp) | LSC length (bp) | SSC length (bp) | IR length (bp) | GC-content | No. of PCGs | GenBank accession no. |
|----------------|-----------|-------|--------------------------------|------------------|-----------------|-----------------|----------------|-------------|-------------|----------------------|
| Crassulaceae   | Crassuloideae |       | Crassula perforata*          | 145,737          | 79,465          | 16,652          | 24,810         | 37.8%       | 85         | MW206794            |
| Kalanchoideae  |           |       | Cotyledon tormentosa*        | 150,049          | 82,250          | 16,995          | 25,402         | 38.2%       | 85         | MW206793            |
|                |           |       | Kalanchoe fedtschenkoi*      | 150,001          | 82,015          | 17,012          | 25,487         | 37.7%       | 85         | MW206796            |
|                |           |       | Kalanchoe tormentosa         | 150,757          | 82,846          | 17,051          | 25,430         | 37.6%       | 85         | MN794319            |
| Sempervivoideae | Acre      |       | Graptotepalum amethystinum*  | 150,365          | 82,009          | 16,764          | 25,396         | 37.9%       | 84         | MW206795            |
|                |           |       | Pachyphytum compactum*       | 149,339          | 81,041          | 16,721          | 25,354         | 37.8%       | 82         | MT860404            |
|                |           |       | Sedum emarginatum           | 149,189          | 81,399          | 16,721          | 25,430         | 37.7%       | 85         | KM281675            |
|                |           |       | Sedum tomentosum            | 149,609          | 81,429          | 16,636          | 25,656         | 37.7%       | 85         | MW206798            |
|                |           |       | Sedum lineare               | 149,257          | 80,963          | 16,484          | 25,823         | 37.9%       | 85         | MT755626            |
|                |           |       | Sedum plumbeiancicola       | 149,397          | 82,121          | 16,670          | 25,783         | 37.7%       | 85         | JX27551             |
|                |           |       | Sedum sarmentosum           | 150,448          | 82,212          | 16,468          | 25,832         | 37.7%       | 85         | MN794333            |
|                |           |       | Aeonium arboreum*           | 150,986          | 82,596          | 16,706          | 25,842         | 37.8%       | 84         | MW206792            |
|                |           |       | Rosularia alpestris         | 151,288          | 82,931          | 16,785          | 25,786         | 37.8%       | 85         | MN794333            |
|                |           |       | Sempervivum           tectorum* | 151,182          | 82,865          | 16,709          | 25,804         | 37.6%       | 85         | MW206799            |
|                |           |       | Hylotelephium           ewersii | 151,699          | 83,253          | 16,838          | 25,804         | 37.7%       | 85         | MN794014            |
|                |           |       | Orostachys           fimbriata* | 151,195          | 82,792          | 16,833          | 25,786         | 37.8%       | 84         | MW206797            |
|                |           |       | Orostachys           japonica | 151,419          | 83,016          | 16,849          | 25,777         | 37.8%       | 85         | MN794320            |
|                |           |       | Phedimus           aizoon | 151,393          | 82,868          | 17,043          | 25,741         | 37.7%       | 85         | MN794321            |
|                |           |       | Phedimus           kamtschaticus | 151,652          | 83,010          | 16,888          | 25,977         | 37.8%       | 85         | MG680403            |
|                |           |       | Rhodiola           integriolata | 151,452          | 82,915          | 17,055          | 25,741         | 37.8%       | 85         | MN794327            |
|                |           |       | Rhodiola           ovatiespala | 151,073          | 82,348          | 17,093          | 25,816         | 37.7%       | 85         | MN794328            |
|                |           |       | Rhodiola           rosea | 151,348          | 82,716          | 17,052          | 25,790         | 37.7%       | 85         | MH410216            |
|                |           |       | Rhodiola           yunnanensis | 151,257          | 82,561          | 17,088          | 25,844         | 37.8%       | 85         | MN794332            |
|                |           |       | Sinocrassula           densirostrata* | 151,773          | 83,123          | 16,904          | 25,873         | 37.7%       | 85         | MW206800            |
|                |           |       | Sinocrassula           indica | 151,755          | 83,159          | 16,888          | 25,854         | 37.7%       | 85         | MN794334            |
|                |           |       | Umbilicus           rupestris | 150,995          | 82,681          | 16,926          | 25,694         | 37.6%       | 85         | MN794335            |
| Haloragaceae s.l. |       |       | Penthorum           chinense | 156,686          | 86,735          | 18,399          | 25,776         | 37.3%       | 84         | JX436155            |
|                |           |       | Glischrocaryon           aureum* | 158,417          | 87,743          | 18,718          | 25,978         | 37.1%       | 83         | MW971555            |
|                |           |       | Glischrocaryon           glandulosum* | 158,146          | 88,123          | 18,743          | 25,640         | 36.8%       | 84         | MW971556            |
|                |           |       | Gonocarpus           micranthus* | 158,655          | 88,165          | 19,000          | 25,745         | 42.8%       | 83         | MW971559            |
|                |           |       | Haloragis           aspera* | 159,395          | 89,207          | 18,482          | 25,853         | 36.7%       | 81         | MW971557            |
|                |           |       | Haloragis           erecta* | 159,414          | 89,043          | 18,555          | 25,908         | 36.7%       | 84         | MW971558            |
|                |           |       | Mynophyllum           spicatum | 158,860          | 88,420          | 18,814          | 25,813         | 36.5%       | 84         | MH191392            |

Species with newly sequenced chloroplast genomes are marked with the asterisks.
LSC, large single-copy region; SSC, small single-copy region; IR, inverted repeat region; PCG, protein-coding gene.
**Pachyphytum compactum** Rose, *Sempervivum tectorum* L., and *Sinocassula densirosulata* (Praeger) A. Berger. In addition, seven complete chloroplast genomes were included as outgroups, i.e., *Myriophyllum spicatum* L., *Penthorum chinense* L., and five new sequenced chloroplast genomes of *Haloragis aspera* Lindl., *Haloragis erecta* (Murray) Oken, *Glischrocaryon aureum* (Lindl.) Orchard, *Glischrocaryon glandulosum* (Orchard) Christenh. and Byng, *Gonocarpus micranthus* Thunb. These seven species belong to the family Haloragaceae *sensu lato*, which are the closest living relatives of Crassulaceae (Jian et al., 2008).

The DNA materials of four species (*Haloragis aspera*, *Haloragis erecta*, *Glischrocaryon aureum* and *Glischrocaryon glandulosum*) were provided by DNA Bank of Royal Botanic Gardens, Kew1. Fresh leaves of the other ten species were collected from the field and preserved with silica gel. The total genomic DNA was extracted using a modified CTAB method (Allen et al., 2006). For each species, one paired-end library with an insertion size of ~350 base pairs (bp) was prepared from the total genomic DNA using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, MA, United States), which was then sequenced using the HiSeq 2500 System (Illumina, Inc., CA, United States) to obtain paired 150-bp reads. Briefly, (i) the genomic DNA was sonicated using the S220 Focused-ultrasonicator (Covaris, MA, United States), (ii) the fragmented DNA was end-repaired, dA-tailed, adapter ligated, and subjected to 10–12 cycles of PCR amplification, and (iii) the quality of each library was assessed using the 2100 Bioanalyzer system (Agilent, CA, United States).

**Chloroplast Genome Assembly and Annotation**

The raw Illumina reads were first filtered to remove paired-end reads if either of the reads contained (i) adapter sequences, (ii) more than 10% of N bases, and (iii) more than 50% of bases with a Phred quality score less than five. The filtered reads were then assembled using NOVOPlasty version 2.7.2 (Dierckxsens et al., 2017), and the complete chloroplast genome of *Sedum sarmentosum* Bunge (Dong et al., 2013) was used as the reference genome. These assemblies were manually inspected using Geneious version 11.0.3 (Kearse et al., 2012).

The assembled chloroplast genomes were annotated using Plann version 1.1 (Huang and Cronk, 2015), and the positions of exons and introns were inspected and adjusted using Sequin version 15.50. In addition, the circular maps of the chloroplast genomes were drawn using OGDRAW version 1.2 (Lohse et al., 2013), and all annotated chloroplast genomes were deposited in GenBank (Sayers et al., 2020).

**Comparative Analysis of Chloroplast Genomes**

For the 26 chloroplast genomes of Crassulaceae, the complete nucleotide sequences were compared using the global alignment algorithm Shuffle-LAGAN (Brudno et al., 2003) as implemented in the program Mvista2 (Frazer et al., 2004). Here, *Rhodiola rosea* L. was chosen as the reference to evaluate gene content variation following Zhao et al. (2020).

To better determine whether any specific pattern of structural variation exists at the family level, the chloroplast genome of *Rhodiola rosea* was used as the representative owing to the highly conserved gene content and gene order within Crassulaceae (see section “Results”), and compared with that of *Penthorum chinense* using progressiveMauve (Darling et al., 2010) as implemented in the software package Mauve version 2.3.1 (Darling et al., 2004). Following Firetti et al. (2017), one of the inverted repeat (IR) regions was manually removed prior to the alignment.

To further identify the hypervariable regions, coding and non-coding regions of the 26 chloroplast genomes were first extracted using PhyloSuite version 1.2.1 (Zhang et al., 2020) and aligned separately using MAFFT version 7.427 (Katoh and Standley, 2013) with default parameters. The nucleotide diversity (π) was then estimated separately for coding and non-coding regions using DnaSP version 6.0 (Rozas et al., 2017).

Moreover, since the size variation of the chloroplast genome may be attributed to the expansion or contraction of the IR region, the boundaries between the IR and single-copy regions were identified using IRscope3 (Amiryousefi et al., 2018) and manually inspected using Geneious.

**Phylogenetic Analysis**

The nucleotide sequences of the chloroplast protein-coding genes were aligned separately using MAFFT and then concatenated into a supermatrix using PhyloSuite. The optimal partitioning scheme and models of DNA sequence evolution were determined using the relaxed hierarchical clustering algorithm (Lanfear et al., 2014) as implemented in PartitionFinder v2.1.1 (Lanfear et al., 2017). Phylogenetic relationships were inferred for 26 species of Crassulaceae (i.e., the 33-taxon supermatrix) using both Bayesian inference (BI) and maximum likelihood (ML) methods. For the BI method, four parallel Markov chain Monte Carlo (MCMC) runs were performed using MrBayes version 3.2.7 (Ronquist et al., 2012). The supermatrix was partitioned based on the optimal scheme determined by PartitionFinder, and the best-fitting substitution model was specified as prior for each partition with model parameters unlinked across partitions. A total of 1,000,000 generations were run with sampling every 500 generations, and the first 25% of samples were discarded as burn-in. Convergence of runs was assumed when the average standard deviation of split frequencies dropped below 0.01. The best-scoring ML tree was inferred using RAxML version 8.2.11 (Stamatakis, 2014) with the GTRGAMMAX model for each partition, and branch support was assessed using the rapid bootstrap algorithm (Stamatakis et al., 2008) with 1,000 replicates.

In order to test the potential effect of uneven taxon sampling from each of the genera, we further subsampled the 26 species of Crassulaceae down to a single species as the representative of each genus. Phylogenetic relationships were then estimated from the 20-taxon supermatrix as described above.

---

1https://dnabank.science.kew.org
2http://genome.lbl.gov/vista/mvista/submit.shtml
3https://irscope.shinyapps.io/trapp/
Positive Selection Analysis

The likelihood ratio test (LRT) and Bayes empirical Bayes (BEB) based on modified branch-site model (Yang et al., 2005; Zhang et al., 2005; Yang and Dos, 2011) were used to identify positively selected genes. Since the topological structures of phylogenetic trees constructed by ML and BI methods were congruent, ML tree was used to positive selection analysis. The amino acid sequences of the sixty-seven common protein-coding genes were aligned using MAFFT and converted into nucleotide alignments using PAL2NAL version 14 (Suyama et al., 2006). The nucleotide alignments were trimmed to obtain the final alignments for positive selection analysis by trimAL version 1.4 (Capellagutiérrez et al., 2009). The branch-site model was performed by codeml program in PAML version 4.9 (Yang, 2007). The branch-site test of positive selection was run with the $\omega$ of foreground lineage fixed to 1 ($\text{fix}_\omega = 1$) for the null hypothesis and estimated ($\text{fix}_\omega = 0$) for the alternative hypothesis. The LRT values at df = 1 were calculated by Chi Square test in PAML, and genes with $p < 0.05$ were treated as candidate positives. Finally, BEB was used to identify those positively selected codon sites.

RESULTS

Characteristics of Chloroplast Genomes in Crassulaceae

After quality control and pre-processing, at least four gigabases (Gb) of whole-genome sequencing data were obtained for each of the nine species (Table 1). These clean reads were assembled...
into high-quality chloroplast genomes using a reference-guided approach, and the resulting coverage ranged from $1,115 \times$ (i.e., Cotyledon tomentosa) to $12,687 \times$ (i.e., Kalanchoe fedtschenkoi). All these newly assembled chloroplast genomes exhibited a typical quadripartite structure, with two IR regions (i.e., IRa and IRb) separating the LSC and SSC regions (Figure 1).

The structure of the chloroplast genome appeared to be largely conserved across the family (Table 1). For each of the 26 Crassulaceae species, the size of the chloroplast genome varied from 145,737 bp (i.e., Crassula perforata) to 151,773 bp (i.e., Sinocrassula densirosulata), and the overall GC-content ranged from 37.6% (i.e., Sempervivum tectorum) to 38.2% (i.e., Cotyledon tomentosa). In addition, the total number of annotated genes in each of these chloroplast genomes ranged from 131 (i.e., Sedum emarginatum Migo) to 134 (i.e., Sedum lineare Thunb.), and all these chloroplast genomes possessed 37 tRNA and four rRNA genes.

Using Rhodiola rosea as the reference, the analysis of mVISTA showed high similarity in gene content and gene order among the 26 chloroplast genomes, and further indicated a fairly high sequence similarity, especially in the coding regions (Supplementary Figure 2). This observation was corroborated by our analysis of nucleotide diversity (Figure 2). The nucleotide diversity in the coding regions ranged from 0 to 0.0794, with an average of 0.0230, which was significantly lower than that in the non-coding regions (0–0.1614, 0.0647; $p$-value < $2.2 \times 10^{-16}$, Welch’s $t$-test). Here, the five coding regions with the highest nucleotide diversity were matK, ycf1, ndhF, rpl22, and rpl32, and the corresponding non-coding regions were trnH-psbA, trnG-trnR, rpl32-trnL, rps16-trnQ, and ccsA-ndhD. Furthermore, no evidence of genomic rearrangement was found in the chloroplast genome of Crassulaceae, when compared with that of Penthorum chinense using progressiveMauve (Supplementary Figure 3).

The boundaries of the LSC, SSC, and IR regions were highly consistent within the family, and no obvious expansion or contraction of the IR region was detected in the 26 chloroplast genomes (Supplementary Figure 4). Here, trnH was shown to be the first gene in the LSC region at the junction between IRa and LSC (i.e., IRa/LSC). At the other end of the LSC region, the junction LSC/IRb was identified as located within the rps19 gene, which gave rise to a truncated copy of the rps19 gene in the IRa region. For both ends of the SSC region, the junctions IRb/SSC and SSC/IRa were found to be located within ndhF and ycf1 gene, respectively. As a consequence, a truncated copy of the ycf1 gene was retained in the IRb region.

**Phylogenetic Relationships of Crassulaceae**

The 33-taxon supermatrix contained a total of 79 genes and 70,905 sites, and the amount of missing data (including gaps and undetermined characters) was 4.5%. Phylogenetic analyses of the 33-taxon supermatrix using ML and BI methods yielded an identical topology (Figure 3), and all relationships were strongly supported by both methods, i.e., $\geq 85$ ML bootstrap percentage (BP) and $\geq 0.99$ Bayesian posterior probability (PP).

All 26 Crassulaceae species formed a monophyletic group, and were divided into three subclades corresponding to the three subfamilies sensu Thiede and Eggli (2007), i.e., Crassuloideae, Kalanchoideae, and Sempervivoideae (Figure 3).
Chang et al. Phylogeny and Chloroplast Genome Structure of Crassulaceae

FIGURE 3 | Phylogenetic inference of 33-taxon supermatrix using maximum likelihood (ML) and Bayesian inference (BI) methods. Branch support was assessed using ML bootstrap percentage (BP) and Bayesian posterior probability (PP), and internal branches with less than 100 BP/1.0 PP are indicated with corresponding values. Species with newly sequenced chloroplast genomes are marked with the asterisks, and clade designations are labeled accordingly.

When rooted with Haloragaceae s.l., Crassuloideae (represented by Crassula perforata) were resolved as sister to Kalanchoeae plus Sempervivoideae. Within Kalanchoeae, the two sampled species of Kalanchoe formed a monophyletic group that was sister to Cotyledon tomentosa. Within Sempervivoideae, our sampled species fell into five clades, i.e., Acre clade, Aeonium clade, Leucosedium clade, Sempervivum clade, and Telephium clade (Figure 3). Here, the Telephium clade was established as sister to the rest of Sempervivoideae, and further split into two lineages. One lineage comprised two sampled species of Phedimus and four sampled species of Rhodiola, which formed two reciprocally monophyletic groups, and the other contained the sampled species of Hylotelephium, Orostachys, Sinocrassula, and Umbilicus. Importantly, Umbilicus rupestris (Salisb.) Dandy was recovered as sister to a clade consisting of Hylotelephium, Orostachys, and Sinocrassula. The Sempervivum clade (represented by Sempervivum tectorum) and the Aeonium clade (represented by Aeonium arboreum) were placed as successive sister lineages to the Leucosedium clade [represented by Rosularia alpestris (Kar. and Kir.) Boriss.] plus the Acre clade. Furthermore, the five sampled species of Sedum in the Acre clade were recovered as a paraphyletic group, with Graptopetalum amethystinum and Pachyphytum compactum nested within them. Importantly, except the non-monophyly of Sedum, the same set of intergeneric relationships was recovered from the 20-taxon supermatrix (Supplementary Figure 5), suggesting that our results should be robust to taxon sampling.

Positive Selection Analysis
A total number of sixty-seven common genes were subjected to positive selection analyses (Table 2). The LRTs with p-value > 0.05 suggested that there was no statistical support for positive selection in any genes (Table 2), although the BEB approach identified fourteen genes (atpB, ndhE, ndhJ, petA, psaC, psaJ, psbB, psbD, psbN, rpoC2, rps15, rps3, rps7, and ycf4) with relatively high posterior probabilities of codon sites.

DISCUSSION
In this study, we report nine newly sequenced complete chloroplast genomes of Crassulaceae. Our comparative analyses indicate that the overall gene organization of the chloroplast genome is highly conserved across all 26 Crassulaceae species.
TABLE 2 | The positive selection test based on the branch-site model.

| Gene name | Null hypothesis | Alternative hypothesis | Significance test |
|-----------|-----------------|------------------------|------------------|
|           | lnL $\omega = 1$ | lnL $\omega > 1$      | BEB             |
| accD      | $-5,981.560$    | $-5,981.396$          | 999.000         | 0.566 |
| atpA      | $-5,342.633$    | $-5,342.633$          | 3.873           | 1.000 |
| atpB      | $-4,756.516$    | $-4,756.516$          | 1.000           | 1.000 |
|           |                 |                       |                 |
| accD      | $-5,981.560$    | $-5,981.396$          | 999.000         | 0.566 |
| atpA      | $-5,342.633$    | $-5,342.633$          | 3.873           | 1.000 |
| atpB      | $-4,756.516$    | $-4,756.516$          | 1.000           | 1.000 |
|           |                 |                       |                 |
| atpE      | $-1,313.375$    | $-1,313.375$          | 1.000           | 1.000 |
| atpF      | $-2,322.492$    | $-2,322.470$          | 436.698         | 0.841 |
| atpH      | $-669.519$      | $-670.763$            | 1.000           | 0.115 |
| atpI      | $-2,397.194$    | $-2,397.194$          | 1.000           | 1.000 |
| ccsA      | $-4,278.565$    | $-4,278.565$          | 3.710           | 1.000 |
| cemA      | $-2,983.494$    | $-2,982.704$          | 999.000         | 0.209 |
| clpP      | $-2,044.145$    | $-2,043.967$          | 71.380          | 1.000 |
| infA      | $-799.954$      | $-799.954$            | 3.322           | 1.000 |
| ndhA      | $-4,381.033$    | $-4,381.036$          | 3.345           | 0.920 |
| ndhB      | $-2,843.138$    | $-2,843.138$          | 3.042           | 1.000 |
| ndhC      | $-1,206.968$    | $-1,206.968$          | 3.034           | 0.997 |
| ndhE      | $-1,161.482$    | $-1,160.256$          | 999.000         | 0.117 |
| ndhG      | $-2,177.605$    | $-2,177.606$          | 3.061           | 1.000 |
| ndhH      | $-4,525.830$    | $-4,525.830$          | 3.179           | 1.000 |
| ndhI      | $-1,904.698$    | $-1,904.698$          | 3.538           | 1.000 |
| ndhJ      | $-1,546.688$    | $-1,544.996$          | 999.000         | 0.075 |
| psaA      | $-6,285.558$    | $-6,285.558$          | 1.000           | 1.000 |
| psaB      | $-6,370.189$    | $-6,370.272$          | 1.000           | 0.689 |
| psaC      | $-650.938$      | $-650.673$            | 26.521          | 0.467 |
| psaD      | $-381.370$      | $-381.370$            | 3.790           | 1.000 |
| psaE      | $-283.178$      | $-283.178$            | 3.396           | 1.000 |
| psaF      | $-283.178$      | $-283.178$            | 3.179           | 1.000 |
| psaI      | $-281.572$      | $-281.572$            | 3.794           | 1.000 |
| psaJ      | $-283.178$      | $-283.178$            | 3.538           | 1.000 |
| psbA      | $-4,124.914$    | $-4,124.914$          | 1.000           | 1.000 |
| psbB      | $-5,011.033$    | $-5,011.033$          | 1.000           | 1.000 |
| psbC      | $-4,124.914$    | $-4,124.914$          | 1.000           | 1.000 |
| psbD      | $-2,835.442$    | $-2,833.539$          | 999.000         | 0.183 |
| psbE      | $-630.873$      | $-630.873$            | 1.000           | 1.000 |
| psbF      | $-234.997$      | $-234.997$            | 1.000           | 1.000 |
| psbG      | $-766.157$      | $-766.157$            | 1.000           | 1.000 |
| psbH      | $-281.572$      | $-281.572$            | 3.794           | 1.000 |
| psbI      | $-283.178$      | $-283.178$            | 3.538           | 1.000 |
| psbK      | $-700.844$      | $-700.824$            | 609.836         | 0.841 |
| psbL      | $-264.296$      | $-264.296$            | 1.000           | 1.000 |
| psbM      | $-329.704$      | $-329.704$            | 1.000           | 1.000 |
| psbN      | $-286.750$      | $-286.002$            | 1.291           | 0.071 |
| psbO      | $-1,410.496$    | $-1,410.496$          | 234.313         | 0.093 |
investigated here. In addition, the results of IRscope analysis reveal no obvious expansion or contraction of the chloroplast IR region. Furthermore, although a recent study has identified a unique 4-kb inversion in the chloroplast genome of the outgroup species Myriophyllum spicatum (Liao et al., 2020), our analyses show a high degree of similarity in chloroplast gene order between Crassulaceae and the outgroup species Penthorum chinense.

Previous studies have demonstrated that the size variation of the angiosperm chloroplast genome is primarily due to variation in the IR region, intergenic region, and gene copy number (Zheng et al., 2017; Amiryousefi et al., 2018; Bedoya et al., 2019). The results of mVISTA analysis suggest that hypervariable regions in the LSC region, such as petA-psbJ, psbM-trnD, psbZ-trnG, rps16-trnQ, and trnE-trnT (Supplementary Figure 2), contribute most to the chloroplast genome size variation within Crassulaceae. Moreover, phylogenetic studies of Crassulaceae have relied primarily on a limited set of chloroplast markers (e.g., matK, rps16, and trnL-trnF). Our comparative analyses, however, have shown that these chloroplast markers appear to be relatively low in nucleotide diversity (Figure 2), which may be partially responsible for the lack of phylogenetic resolution within Sempervivoideae (Mort et al., 2001; Messerschmid et al., 2020). Thus, to achieve better phylogenetic resolution, future studies of Crassulaceae should focus on molecular markers from more variable regions of the chloroplast genome, such as ccsA-ndhD, rps16-trnQ, rpl32-trnL, and trnH-psbA (Prince, 2015).

With increased taxon sampling of key lineages of Crassulaceae, our phylogenetic analyses of chloroplast genome sequences have substantially improved the phylogenetic resolution, and provided robust inference of the infrafamilial relationships. Among the three subfamilies sensu Thiede and Eggli (2007), our phylogenetic analyses have confirmed the monophyly of Kalanchoideae and of Sempervivoideae, and suggested that Crassuloidae are sister to Kalanchoideae plus Sempervivoideae, corroborating previous studies based on DNA sequences and restriction site variation (e.g., van Ham and Hart, 1998; Mort et al., 2001, 2009; Bruyns et al., 2019; Folk et al., 2019). In addition, the sister relationship between Kalanchoideae and Sempervivoideae is supported by two putative morphological synapomorphies, i.e., leaves with a single apical or subapical hydathode and seeds with costate testa (Thiede and Eggli, 2007).

Numerous studies have attempted to identify major lineages within Sempervivoideae (e.g., van Ham and Hart, 1998; Mort et al., 2001; Mayuzumi and Ohba, 2004; Nikulin et al., 2016; Folk et al., 2019; Messerschmid et al., 2020), but relationships among these major lineages remain uncertain (Supplementary Figure 1). Our analyses have revealed five clades that are successively sister lineages, i.e., Telephium clade, Sempervivum clade, Aeonium clade, Leucosedum clade, and Acre clade. These clades correspond to five out of the six major clades sensu Messerschmid et al. (2020). Of these six major clades, the lone exception here is the Petrosedum clade sensu Messerschmid et al. (2020), for which no chloroplast genome sequence is currently available. The Telephium clade identified here is equivalent to

| Gene name | Null hypothesis | Alternative hypothesis | Significance test | P-value |
|-----------|----------------|-----------------------|------------------|---------|
|           | lnL ω = 1      | lnL ω > 1             | BEB              |         |
| rpl23     | −514.597       | −514.594              | 5.914            | 0.943   |
| rpl32     | −768.421       | −768.421              | 3.262            | 1.000   |
| rpl33     | −770.969       | −770.721              | 999.000          | 0.480   |
| rpl36     | −358.642       | −358.642              | 6.466            | 1.000   |
| rpoA      | −4,186.019     | −4,185.993            | 999.000          | 0.823   |
| rpoB      | −12,323.016    | −12,323.016           | 2.583            | 1.000   |
| rpoC2     | −18,495.025    | −18,494.194           | 4.364            | 0.198   |
| rps11     | −1,415.271     | −1,415.250            | 48.778           | 0.841   |
| rps12     | −765.506       | −765.397              | 113.044          | 0.590   |
| rps14     | −989.959       | −989.959              | 1.000            | 1.000   |
| rps15     | −1,369.457     | −1,369.056            | 8.032            | 0.370   |
| rps18     | −849.939       | −849.939              | 1.000            | 1.000   |
| rps19     | −1,186.279     | −1,186.279            | 1.979            | 1.000   |
| rps2      | −2,405.932     | −2,405.934            | 1.000            | 1.000   |
| rps3      | −2,700.278     | −2,700.278            | 1.000            | 1.000   |
| rps4      | −1,960.498     | −1,960.497            | 1.000            | 1.000   |
| rps7      | −786.610       | −786.345              | 49.908           | 0.467   |
| rps8      | −1,693.306     | −1,693.306            | 1.000            | 1.000   |
| ycf3      | −1,302.337     | −1,302.337            | 1.000            | 1.000   |
| ycf4      | −2,054.432     | −2,054.432            | 1.000            | 1.000   |

Bold types are genes with positively selected sites identified by BEB with relatively high posterior probabilities.
tribes Telephieae plus Umbiliceae sensu Thiede and Eggli (2007). Although it was first proposed by van Ham and Hart (1998), the monophyly of the Telephium clade has only recently been confirmed (i.e., >85 BP) by phylogenetic analyses of chloroplast genome sequences (Kim and Kim, 2020) and 301 low-copy nuclear genes (Folk et al., 2019). Our results add further evidence to support the recognition of the Telephium clade. In addition, Umbilicus rupestris is strongly supported by our phylogenetic analyses as sister to the tribe Telephieae (i.e., our sampled species of Hylotelephium, Orostachys, and Sinocrassula), thus corroborating recent findings (Folk et al., 2019; Kim and Kim, 2020; Zhao et al., 2020) and highlighting the paraphyly of the tribe Umbiliceae (i.e., our sampled species of Phedimus, Rhodiola, and Umbilicus). Furthermore, the Sempervivum and Aeonium clades identified here correspond to tribes Semperviveae and Aeonieae sensu Thiede and Eggli (2007), respectively, and the Leucosedum and Acre clades identified here together constitute tribe Sedeae sensu Thiede and Eggli (2007).

It has been indicated that the genes with positive selection played key parts in the adaptation to diverse environments (Moseley et al., 2018; Gao et al., 2019b; Li et al., 2020b). However, no positive selection was statistically supported among sixty-seven chloroplast protein-coding genes in sampled Crassulaceae. This result indicated these protein-coding genes are under strong structural and functional constraints. The absence of positive selection in most protein-coding genes of chloroplast genome was also found in Euonymus (Li et al., 2021) and Quercus (Yin et al., 2018).

In conclusion, by expanding the number of informative molecular characters, we have further improved the resolution of the phylogenetic relationships among major lineages within Crassulaceae, which will facilitate the identification of non-molecular synapomorphies. However, additional sampling of key lineages (e.g., Petrosedum clade) is required to fully resolve the infrafamilial relationships. Furthermore, the boundaries of some of the traditional genera in Crassulaceae remain poorly defined. For example, Sedum, the largest genus with approximately 470 species, has been shown to be highly polyphyletic, and its intrageneric relationships remain largely unresolved (Messerschmid et al., 2020). Thus, chloroplast phylogenomics will continue to enhance our understanding of the evolutionary history of Crassulaceae in the future.

DATA AVAILABILITY STATEMENT
All annotated chloroplast genomes have been deposited in GenBank (https://www.ncbi.nlm.nih.gov/genbank/), and accession numbers are provided in Table 1.

AUTHOR CONTRIBUTIONS
JL, ZX, and XX designed research. HC, LZ, and HX performed research and analyzed data. HC, JL, ZX, and XX wrote the manuscript. All authors reviewed and revised the manuscript.

FUNDING
This work was supported by the National Key Research and Development Program of China (2017YFC0505203), the National Natural Science Foundation of China (31770566, 31600172), and Biodiversity Survey, Observation and Assessment Program of Ministry of Ecology and Environment of China.

ACKNOWLEDGMENTS
We would like to thank Quanjun Hu, Minghui Kang, Tao Ma, Deyan Wang, and Xuchen Yang for technical assistance and valuable discussions. We would also like to thank DNA Bank of Royal Botanic Gardens for providing DNA materials of Haloragaceae.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.631884/full#supplementary-material

REFERENCES
Acevedo-Rosas, R., Cameron, K., Sosa, V., and Pelli, S. (2004). A molecular phylogenetic study of graptopetalum (Crassulaceae) based on ETS, ITS, rpl16, and trnL-F nucleotide sequences. Am. J. Bot. 91, 1099–1104. doi: 10.10372/ajb.91,7.1099
Allen, G. C., Flores-Vergara, M. A., Krasnyanski, S., Kumar, S., and Thompson, W. F. (2006). A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nat. Protoc. 1, 2320–2325. doi: 10.1038/nprot.2006.384
Amiryousefi, A., Hyvönen, J., and Poczaí, P. (2018). IRscope: an online program to visualize the junction sites of chloroplast genomes. Bioinformatics 34, 3030–3031. doi: 10.1093/bioinformatics/bty220
APG IV. (2016). An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. Bot. J. Linn. Soc. 181, 1–20. doi: 10.1111/bij.12385
Bedoya, A. M., Ruhfeld, B. R., Philbrick, C. T., Madriñán, S., Bove, C. P., Mesterházy, A., et al. (2019). Plastid genomes of five species of riverweeds (Podostemaceae): structural organization and comparative analysis in Malpighiales. Front. Plant Sci. 10:1035. doi: 10.3389/fpls.2019.01035
Berger, A. (1930). “Crassulaceae”, in Die Natürlichen Pflanzenfamilien, Band 18a, eds A. Engler and K. Prantl (Leipzig: Wilhelm Engelmann), 352–483.
Bruino, M., Malde, S., Polakow, A., Do, C. B., Couronne, O., Dubchak, I., et al. (2003). Glocal alignment: finding rearrangements during alignment. Bioinformatics 19, i54–i62.
Bruyns, P. V., Hanáˇcek, P., and Klak, C. (2019). Crassula, insights into an old, arid-adapted group of southern African leaf-succulents. Mol. Phylogenet. Evol. 131, 35–47. doi: 10.1016/j.ympev.2018.10.045
Capellagutiérrez, S., Sillamartínez, J. M., and Gabaldón, T. (2009). Trimal: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973. doi: 10.1093/bioinformatics/btp348
Carrillo-Reyes, P., Sosa, V., and Mort, M. E. (2008). Thompsonella and the “Echeveria group” (Crassulaceae): phylogenetic relationships based on molecular and morphological characters. Taxon 57, 863–874. doi: 10.1002/tax.573015
that Amborella is not a basal angiosperm. Mol. Biol. Evol. 20, 1499–1505. doi: 10.1093/molbev/msg159

Goremykin, V. V., Hirsch-Ernst, K. L., Wolff, S., and Hellwig, F. H. (2004). The chloroplast genome of Nymphaea alba: whole-genome analyses and the problem of identifying the most basal angiosperm. Mol. Biol. Evol. 21, 1445–1454. doi: 10.1093/molbev/msi147

Gruzev, E. V., Kadnikov, V. V., Beletsky, A. V., Mardanov, A. V., and Ravin, N. V. (2019). Extensive plastome reduction and loss of photosynthesis genes in Diplhypaea coccinea, a holoparasitic plant of the family Orobanchaceae. PeerJ. 7:e7830. doi: 10.7717/peerj.7830

Hu, S.-L., Sablok, G., Wang, B., Qu, D., Barbaro, E., Viola, R., et al. (2015). Plastome organization and evolution of chloroplast genes in Cardamine species adapted to contrasting habitats. BMC Genom. 16:306. doi: 10.1186/s12864-015-1498-0

Hu, Y.-H., Zhou, K. E., and Zhao, P. (2016). Completion of the chloroplast genomes of five Chinese juglans and their contribution to chloroplast phylogeny. Front. Plant Sci. 7:1955. doi: 10.3389/fpls.2016.01955

Huang, D. I., and Cronk, Q. C. B. (2015). Plann: a command-line application for annotating plastome sequences. PeerJ. 3:150026. doi: 10.7732/appps.150026

Huo, Y.-L., Gao, L.-M., Liu, B.-J., Yang, Y.-Y., Kong, S. P., Sun, Y.-Q., et al. (2019). Complete chloroplast genome sequences of four Allium species: comparative and phylogenetic analyses. Sci. Rep. 9:12250.

Jian, S., Soltis, P. S., Gitzenmanner, D. A., Moore, M. J., Li, R., Hendry, T. A., et al. (2008). Resolving an ancient, rapid radiation in Saxifragales. Syst. Biol. 57, 38–57. doi: 10.1080/10557211003448871

Jorgensen, T. H., and Frydenberg, J. (1999). Diversification in insular plants: inferring the phylogenetic relationship in Aeonium (Crassulaceae) using ITS sequences of nuclear ribosomal DNA. Nord. J. Bot. 19, 613–621. doi: 10.1111/j.1756-1051.1999.tb01150.x

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. doi: 10.1093/molbev/mst010

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kim, S.-H., and Kim, S.-C. (2020). Plastid genome of stonecrop hydropetalum verticillatum (Sedoideae; Crassulaceae): insight into structure and phylogenetic position. Mitochondrial DNA B Resour. 5, 2729–2731. doi: 10.1080/23802359.2020.1788464

Klein, J. T., and Kaderetz, J. W. (2015). Phylogeny, biogeography, and evolution of edaphic association in the European orophytes sempervivum and jovibarba (Crassulaceae). Int. J. Plant Sci. 176, 64–71. doi: 10.1086/677948

Lanfear, R., Calcott, B., Kainer, D., Mayer, C., and Stamatakis, A. (2014). Selecting optimal partitioning schemes for phylogenetic datasets. BMC Evol. Biol. 14:82. doi: 10.1186/1471-2148-14-82

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., and Calcott, B. (2017). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34, 772–773.

Li, H.-W., Liu, B., Davis, C. C., and Yang, Y. (2020a). Plastome phylogenomics, systematics, and divergence time estimation of the Beilschmiedia group (Lauraceae). Mol. Phylogenet. Evol. 151:106901. doi: 10.1016/j.ympev.2020.106901

Li, C.-Y., Zhao, Y.-L., Xu, Z.-G., Yang, G.-Y., and Peng, J. (2020b). Initial plastome phylogenomics, for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 33, 7:1955. doi: 10.3389/fpls.2020.593984

Liao, Y.-Y., Liu, Y., Liu, X., Liu, T.-F., Mbichi, R. W., Wan, T., et al. (2020). The complete chloroplast genome of Myriophyllum spicatum reveals a 4-kb
inversion and new insights regarding plastome evolution in Haloragaceae. *Ecol. Evol.* 10, 3090–3102. doi: 10.1002/ece3.6125

Lohse, M., Drechsel, O., Kahlau, S., and Bock, R. (2013). OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* 41, W575–W581

Magallón, S., Sánchez-Reyes, L. L., and Gómez-Acevedo, S. L. (2019). Thirty clues to the exceptional diversification of flowering plants. *Ann. Bot.* 123, 491–503. doi: 10.1093/aob/mcy182

Mayuzumi, S., and Ohba, H. (2004). The phylogenetic position of eastern Asian sedoideae (Crassulaceae) inferred from chloroplast and nuclear DNA sequences. *Syst. Bot.* 29, 587–598. doi: 10.1600/03644041744329

Mes, T. H. M., Wijers, G.-J., and Hart, H. (1997). Phylogenetic relationships in Monanthes (Crassulaceae) based on morphological, chloroplast and nuclear DNA variation. *J. Evol. Biol.* 10, 193–216. doi: 10.1046/j.1420-9101.1997.1002019.x

Messerenschmid, T. F. E., Klein, J. T., Kadereit, G., and Kadereit, J. W. (2013). *Lohse, M., Drechsel, O., Kahlau, S., and Bock, R. (2013). OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Res.* 41, W575–W581

Messerenschmid, T. F. E., Klein, J. T., Kadereit, G., and Kadereit, J. W. (2013). *Lohse, M., Drechsel, O., Kahlau, S., and Bock, R. (2013). OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Res.* 41, W575–W581

Nikulin, V. Y., Gontcharova, S. B., Stephenson, R., and Goncharov, A. A. (2002). Phylogenetics and evolution of the macaronesian clade of Crassulaceae inferred from nuclear and chloroplast DNA sequence data. *Am. J. Bot.* 92, 1170–1176. doi: 10.3732/ajb.92.7.1170

Mort, M. E., Levsen, N., Randle, C. P., Van Jaarsveld, E., and Palmer, A. (2005). Phylogenetics and diversification of Cotyledon (Crassulaceae) inferred from cpDNA data sequence data. *Ann. Bot.* 91, 271–288. doi: 10.1006/anbo.2004.1364

Moseley, R. C., Mewalal, R., Motta, F., Tuskan, G. A., Haase, S., and Yang, X. (2016). *Messerschmid, T. F. E., Klein, J. T., Kadereit, G., and Kadereit, J. W. (2013).* Lohse, M., Drechsel, O., Kahlau, S., and Bock, R. (2013). OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* 41, W575–W581

Mort, M. E., Randle, C. P., Burgos, P., Smith, G., Jaarsveld, E., and Hopper, S. D. (2005). Phylogenetic relationships and evolution of Crassulaceae inferred from mtDNA sequence data. *Am. J. Bot.* 92, 76–91. doi: 10.3732/ajb.92.7.125

Nikulin, V. Y., Gontcharova, S. B., Stephenson, R., and Goncharov, A. A. (2016). *Messerschmid, T. F. E., Klein, J. T., Kadereit, G., and Kadereit, J. W. (2013).* Lohse, M., Drechsel, O., Kahlau, S., and Bock, R. (2013). OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* 41, W575–W581

Pringle, L. M. (2015). Plastid primers for angiosperm phylogenetics and phylogeography. *Appl. Plant Sci.* 3:1400085. doi: 10.3732/apps.1400085

Prince, L. M. (2015). Plastid primers for angiosperm phylogenetics and phylogeography. *Appl. Plant Sci.* 3:1400085. doi: 10.3732/apps.1400085

Quezada, I. M., and Gianoli, E. (2011). Crassulacean acid metabolism. *Flora.* 206, 348–355. doi: 10.1016/j.flora.2011.03.025

Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302. doi: 10.1093/molbev/msx248

RaxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu333

Stamatakis, A. (2014). Phylogenetic analysis and character evolution of Rhodiola (Crassulaceae) based on comparative analysis of the complete chloroplast genome of Quercus aquifolii with other oak species. *Int. J. Mol. Sci.* 19:1042. doi: 10.3390/ijms19101042

Thiede, J., and Eggl, U. (2007). “Crassulaceae,” in *Vascular Plants, ed. K. Kubitzki (Berlin: Springer), 83–118.

Twyford, A. D., and Ness, R. W. (2017). Strategies for complete plastome genome sequencing. *Ann. Bot.* 119, 858–868. doi: 10.1093/aob/mcx126

van Ham, R., and Hart, H. (1998). Phylogenetic relationships in the Crassulaceae inferred from chloroplast DNA restriction-site variation. *Ann. Bot.* 85, 123–134. doi: 10.1006/anbo.1998.0246

Vázquez-Cotero, C., Sosa, V., and Carrillo-Reyes, P. (2017). Phylogenetic position of echeveria heterosepala (Crassulaceae): a rare species with diagnostic characters of PachypHYTUM. *Bot. Soc. 95, 515–526. doi: 10.17129/botsci.1190

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088

Yang, Z., Wang, W. S. W., and Nielsen, R. (2005). Bayes empirical bayes inference of amino acid sites under positive selection. *Mol. Biol. Evol.* 22, 1107–1118. doi: 10.1093/molbev/msi097

Yan, Z.-H., and Dos, R.-M. (2011). Statistical properties of the branch-site test of positive selection. *Mol. Biol. Evol.* 28, 1217–1228. doi: 10.1093/molbev/msq303

Yin, K., Zhang, Y., Li, Y., and Du, F. (2018). Different natural selection pressures on the atpF gene in evergreen sclerophyllous and Deciduous oak species: evidence from comparative analysis of the complete chloroplast genome of Quercus aquifolii with other oak species. *Int. J. Mol. Sci.* 19:1042. doi: 10.3390/ijms19101042

Yost, J. M., Bontrager, M., McCabe, S. W., Burton, D., Simpson, M. G., Kay, K. M., et al. (2013). Phylogenetic relationships and evolution in Dudleya (Crassulaceae). *Syst. Bot.* 38, 1096–1104. doi: 10.1600/03634413x47760

Zhang, D., Yuan, T., and Zhang, Y.-H. (2018). Conservation and innovation: plastome evolution during rapid radiation of Rhodolia (Crassulaceae) based on nuclear ribosomal ITS and plastid trnL-F and psbA-trnH sequences. *Syst. Bot.* 43, 441–451. doi: 10.1600/03634413x480753

Zhao, D.-N., Ren, Y., and Zhang, J.-Q. (2020). Conservation and innovation: plastome evolution during rapid radiation of Rhodolia on the Qinghai-Tibetan Plateau. *Mol. Phylogenet. Evol.* 144:106713. doi: 10.1016/j.ympev.2019.05.030

Zhao, D.-N., Ren, Y., and Zhang, J.-Q. (2020). Conservation and innovation: plastome evolution during rapid radiation of Rhodolia on the Qinghai-Tibetan Plateau. *Mol. Phylogenet. Evol.* 144:106713. doi: 10.1016/j.ympev.2019.05.030
Zheng, X.-M., Wang, J., Feng, L., Liu, S., Pang, H., Qi, L., et al. (2017). Inferring the evolutionary mechanism of the chloroplast genome size by comparing whole-chloroplast genome sequences in seed plants. Sci. Rep. 7:1555.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.